Rank Equalities Related to Generalized Inverses of Matrices and Their Applications

Yongge Tian

ACKNOWLEDGEMENTS

Many thanks go to my thesis advisor George P. H. Styan for his guidance, invaluable discussions and thorough understanding during the preparation of this thesis.

Sincere thanks also go to Professor Ronald J. Stern for his advice and encouragement.

ABSTRACT

This paper is divided into two parts. In the first part, we develop a general method for expressing ranks of matrix expressions that involve Moore-Penrose inverses, group inverses, Drazin inverses, as well as weighted Moore-Penrose inverses of matrices. Through this method we establish a variety of valuable rank equalities related to generalized inverses of matrices mentioned above. Using them, we characterize many matrix equalities in the theory of generalized inverses of matrices and their applications.

In the second part, we consider maximal and minimal possible ranks of matrix expressions that involve variant matrices, the fundamental work is concerning extreme ranks of the two linear matrix expressions $A - BXC$ and $A - B_1X_1C_1 - B_2X_2C_2$. As applications, we present a wide range of their consequences and applications in matrix theory.

AMS subject classifications: 15A03, 15A09, 15A24.

Key words: Rank, inner inverse, Moore-Penrose inverse, group inverse, Drazin inverse, reverse order law, EP matrix, block matrix, Schur complement, matrix expression, matrix equation, solution.

Contents

Chapter 1

Introduction and preliminaries

This is a comprehensive work on ranks of matrix expressions involving Moore-Penrose inverses, group inverses, Drazin inverses, as well as weighted Moore-Penrose inverses. In the theory of generalized inverses of matrices and their applications, there are numerous matrix expressions and equalities that involve these three kinds of generalized inverses of matrices. Now we propose such a problem: Let $p(A_1^\dagger, \cdots, A_k^\dagger)$ and $q(B_1^\dagger, \cdots, B_l^\dagger)$ be two matrix expressions involving Moore-Penrose inverses of matrices. Then determine necessary and sufficient conditions such that $p(A_1^\dagger, \cdots, A_k^\dagger) = q(B_1^\dagger, \cdots, B_l^\dagger)$ holds. A seemingly trivial condition for this equality to hold is apparently

$$\operatorname{rank}[\, p(A_1^\dagger, \cdots, A_k^\dagger) - q(B_1^\dagger, \cdots, B_l^\dagger)\,] = 0. \tag{1.1}$$

However, if we can reasonably find a formula for expressing the rank of the left-hand side of (1.1), then we can derive immediately from (1.1) nontrivial conditions for

$$p(A_1^\dagger, \cdots, A_k^\dagger) = q(B_1^\dagger, \cdots, B_l^\dagger)$$

to hold. This work has a far-reaching influence to many problems in the theory of generalized inverses of matrices and their applications. This consideration motivates us to make a thorough investigation to this work. In fact, the author has successfully used this idea to establish necessary and sufficient conditions such that $(ABC)^\dagger = C^\dagger B^\dagger A^\dagger$, $(ABC)^\dagger = (BC)^\dagger B(AB)^\dagger$, and $(A_1A_2\cdots A_k)^\dagger = A_k^\dagger \cdots A_2^\dagger A_1^\dagger$ (cf. [133], [135]). But the methods used in those papers are somewhat restricted and not applicable to various kinds of matrix expressions. In this paper, we shall develop a general and complete method for establishing rank equalities for matrix expressions involving Moore-Penrose inverses, group inverses, Drazin inverses, as well as weighted Moore-Penrose inverses of matrices. Using these rank formulas, we shall characterize various equalities for generalized inverses of matrices, and then present their applications in the theory of generalized inverses of matrices.

The matrices considered in the this paper are mainly over the complex number field $\mathcal{C}$. Let $A \in \mathcal{C}^{m\times n}$. We use A^*, $r(A)$ and $R(A)$ to stand for the conjugate transpose, the rank and the range (column space) of A, respectively.

It is well known that the Moore-Penrose inverse of matrix A is defined to be the unique solution X of the following four Penrose equations

$$(1) \quad AXA = A, \qquad (2) \quad XAX = X, \qquad (3) \quad (AX)^* = AX, \qquad (4) \quad (XA)^* = XA,$$

and is often denoted by $X = A^\dagger$. In addition, a matrix X that satisfies the first equation above is called an inner inverse of A, and often denoted by A^-. A matrix X that satisfies the second equation above is called an outer inverse of A, and often denoted by $A^{(2)}$. For simplicity, we use E_A and F_A to stand for the two projectors

$$E_A = I - AA^\dagger \qquad \text{and} \qquad F_A = I - A^\dagger A$$

induced by A. As to various basic properties concerning Moore-Penrose inverses of matrices, see, e.g., Ben-Israel and Greville [16] (1980), Campbell and Meyer [21] (1991), Rao and Mitra [118] (1971).

Let $A \in \mathcal{C}^{m \times m}$ be given with $\operatorname{Ind} A = k$, the smallest positive integer such that $r(A^{k+1}) = r(A^k)$. The Drazin inverse of matrix A is defined to be the unique solution X of the following three equations

$$(1) \quad A^k X A = A^k, \qquad (2) \quad XAX = X, \qquad (3) \quad AX = XA,$$

and is often denoted by $X = A^D$. In particular, when $\operatorname{Ind} A = 1$, the Drazin inverse of matrix A is call the group inverse of A, and is often denoted by $A^{\#}$.

Let $A \in \mathcal{C}^{m \times n}$. The weighted Moore-Penrose inverse of $A \in \mathcal{C}^{m \times n}$ with respect to the two positive definite matrices $M \in \mathcal{C}^{m \times m}$ and $N \in \mathcal{C}^{n \times n}$ is defined to be the unique solution of the following four matrix equations

$$(1) \quad AXA = A, \qquad (2) \quad XAX = X, \qquad (3) \quad (MAX)^* = MAX, , \qquad (4) \quad (NXA)^* = NXA,$$

and this X is often denoted by $X = A^{\dagger}_{M,N}$. In particular, when $M = I_m$ and $N = I_n$, $A^{\dagger}_{M,N}$ is the conventional Moore-Penrose inverse $A^{\dagger}$ of A. Various basic properties concerning Drazin inverses, group inverses and Weighted Moore-Penrose inverses of matrices can be found in Ben-Israel and Greville [16], Campbell and Meyer [21], Rao and Mitra [118].

It is well known that generalized inverses of matrices are a powerful tool for establishing various rank equalities for matrices. Two seminal references are the paper [82] by Marsaglia and Styan (1974) and the paper [95] by Meyer (1973). In those two papers, some fundamental rank equalities and inequalities related to generalized inverses of block matrices were established and a variety of consequences and applications of these rank equalities and inequalities were considered. Since then, the main results in those two papers have widely been applied to dealing with various problems in the theory of generalized inverses of matrices and its applications. To some extent, this paper could be regarded as a summary and extension of all work related to those two remarkable papers.

We next list some key results in those papers, which will be intensively applied in this monograph.

Lemma 1.1 (Marsaglia and Styan [82], Meyer [95]). *Let $A \in \mathcal{C}^{m \times n}$, $B \in \mathcal{C}^{m \times k}$, $C \in \mathcal{C}^{l \times n}$ and $D \in \mathcal{C}^{l \times k}$. Then*

$$r[\, A, \ B\,] = r(A) + r(\, B - AA^{\dagger}B\,) = r(B) + r(\, A - BB^{\dagger}A\,), \tag{1.2}$$

$$r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r(\, C - CA^{\dagger}A\,) = r(C) + r(\, A - AC^{\dagger}C\,), \tag{1.3}$$

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(B) + r(C) + r(E_BAF_C) = r(B) + r(C) + r[\,(\, I_m - BB^{\dagger}\,)A(\, I_n - C^{\dagger}C\,)\,], \tag{1.4}$$

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r(A) + r\begin{bmatrix} 0 & E_AB \\ CF_A & S_A \end{bmatrix}, \tag{1.5}$$

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A \\ C \end{bmatrix} + r[\, A, \ B\,] - r(A) + r[J(D)], \tag{1.6}$$

where $S_A = D - CA^{\dagger}B$ is the Schur complement of A in $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, and

$$J(D) = [\, I - (CF_A)(CF_A)^{\dagger}\,]S_A[\, I - (E_AB)^{\dagger}(E_AB)\,],$$

called the rank complement of D in M. In particular, if $R(B) \subseteq R(A)$ and $R(C^*) \subseteq R(A^*)$, then

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r(A) + r(\, D - CA^{\dagger}B\,). \tag{1.7}$$

The six rank equalities in (1.2)—(1.7) *are also true when replacing $A^{\dagger}$ by any inner inverse A^- of A.*

Lemma 1.2[82]. *Let $A \in \mathcal{C}^{m \times n}$, $B \in \mathcal{C}^{m \times k}$, $C \in \mathcal{C}^{l \times n}$ and $D \in \mathcal{C}^{l \times k}$. Then*

(a) $r[\, A, \ B\,] = r(A) + r(B) \Leftrightarrow R(A) \cap R(B) = \{0\} \Leftrightarrow R[(E_AB)^*] = R(B^*) \Leftrightarrow R[(E_BA)^*] = R(A^*)$.

(b) $r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r(C) \Leftrightarrow R(A^*) \cap R(C^*) = \{0\} \Leftrightarrow R(CF_A) = R(C) \Leftrightarrow R(AF_C) = R(A)$.

(c) $r[\,A,\ B\,] = r(A) \Leftrightarrow R(B) \subseteq R(A) \Leftrightarrow E_A B = 0.$

(d) $r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) \Leftrightarrow R(C^*) \subseteq R(A^*) \Leftrightarrow CF_A = 0.$

(e) $r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(A) + r(B) + r(C) \Leftrightarrow R(A) \cap R(B) = \{0\}$ *and* $R(A^*) \cap R(C^*) = \{0\}.$

(f) $r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r(A) \Leftrightarrow R(B) \subseteq R(A)$ *and* $R(C^*) \subseteq R(A^*)$ *and* $D = CA^{\dagger}B.$

Lemma 1.3[82](*Rank cancellation rules*). *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given, and suppose that*

$$R(AQ) = R(A) \quad \text{and} \quad R[(PA)^*] = R(A^*).$$

Then

$$r[\,AQ,\ B\,] = r[\,A,\ B\,], \qquad r\begin{bmatrix} PA \\ C \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix}. \tag{1.8}$$

In particular,

$$r[\,AA^*,\ B\,] = r[\,A,\ B\,], \qquad r\begin{bmatrix} A^*A \\ C \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix}. \tag{1.9}$$

Lemma 1.4[82]. *Let $A,\ B \in \mathcal{C}^{m\times n}$. Then*

(a) $r(\,A \pm B\,) \geq r\begin{bmatrix} A \\ B \end{bmatrix} + r[\,A,\ B\,] - r(A) - r(B).$

(b) *If $R(A) \cap R(B) = \{0\}$, then $r(\,A + B\,) = r\begin{bmatrix} A \\ B \end{bmatrix}$.*

(c) *If $R(A^*) \cap R(B^*) = \{0\}$, then $r(\,A + B\,) = r[\,A,\ B\,]$.*

(d) $r(\,A+B\,) = r(A)+r(B) \Leftrightarrow r(\,A-B\,) = r(A)+r(B) \Leftrightarrow R(A)\cap R(B) = \{0\}$, *and* $R(A^*)\cap R(B^*) = \{0\}$.

In addition, we shall also use in the sequel the following several basic rank formulas, which are either well known or easy to prove.

Lemma 1.5. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times m}$ and $N \in \mathcal{C}^{m\times m}$. Then*

$$r(\,A - ABA\,) = r(A) + r(\,I_n - BA\,) - n = r(A) + r(\,I_m - AB\,) - m, \tag{1.10}$$

$$r(\,N \pm N^{k+1}\,) = r(N) + r(\,I_n \pm N^k\,) - m, \quad \text{for all } k \geq 1, \tag{1.11}$$

$$r(\,I_m - N^2\,) = r(\,I_m + N\,) + r(\,I_m - N\,) - m. \tag{1.12}$$

Lemma 1.6. *Let $A,\ B \in \mathcal{C}^{m\times n}$. Then*

$$r\begin{bmatrix} A & B \\ B & A \end{bmatrix} = r(\,A + B\,) + r(\,A - B\,). \tag{1.13}$$

Proof. Follows from the following decomposition

$$\frac{1}{2}\begin{bmatrix} I_m & I_m \\ I_m & -I_m \end{bmatrix}\begin{bmatrix} A & B \\ B & A \end{bmatrix}\begin{bmatrix} I_n & I_n \\ I_n & -I_n \end{bmatrix} = \begin{bmatrix} A+B & 0 \\ 0 & A-B \end{bmatrix}.$$

Lemma 1.7. *Let $A \in \mathcal{C}^{m\times m}$. Then*

$$r(\,A - A^{2k+1}\,) = r(\,A + A^k\,) + r(\,A - A^k\,) - r(A) = r(A) + r(\,I_m + A^{k-1}\,) + r(\,I_m - A^{k-1}\,) - 2m. \tag{1.14}$$

In particular (Anderson and Styan [1])

$$r(\,A - A^3\,) = r(\,A + A^2\,) + r(\,A - A^2\,) - r(A) = r(A) + r(\,I_m + A\,) + r(\,I_m - A\,) - 2m. \tag{1.15}$$

Proof. Replace B in (1.13) by A^k and simplify to yield (1.14). $\square$

Lemma 1.8. *Let $A \in \mathcal{C}^{m\times m}$ and $\lambda_1, \lambda_2, \cdots, \lambda_k \in \mathcal{C}$ with $\lambda_i \neq \lambda_j$ for $i \neq j$. Then for any positive integer $t_1, t_2, \cdots, t_k$, the following rank equality*

$$\begin{aligned} & r[\,(\lambda_1 I - A)^{t_1}(\lambda_2 I - A)^{t_2}\,\cdots\,(\lambda_k I - A)^{t_k}\,] \\ & = r[\,(\,\lambda_1 I - A\,)^{t_1}] + r[(\,\lambda_1 I - A\,)^{t_2}] + \cdots + r[(\,\lambda_k I - A\,)^{t_k}\,] - (k-1)m. \end{aligned} \tag{1.16}$$

always holds. This result can also alternatively be stated that for any two polynomials $p(x)$ and $q(x)$ without common roots, there is

$$r[\,p(A)q(A)\,] = r[\,p(A)\,] + r[\,q(A)\,] - m. \tag{1.17}$$

This paper is divided into two parts with 31 chapters. They organize as follows.

In Chapter 2, we establish several universal rank formulas for matrix expressions that involve Moore-Penrose inverses of matrices. These rank formulas will serve as a basic tool for developing the content in the subsequent chapters.

In Chapter 3, we present a set of rank formulas related to sums, differences and products of idempotent matrices. Based on them, we shall reveal a series of new and nontrivial properties for idempotent matrices.

In Chapter 4, we extend the results in Chapter 3 to some matrix expressions that involve both idempotent matrices and general matrices. In addition, we shall also establish a group of new rank formulas related to involutory matrices and then consider their consequences.

In Chapter 5, we establish a set of rank formulas related to outer inverses of a matrix. Some of them will be applied in the subsequent chapters.

In Chapter 6, we examine various relationships between a matrix and its Moore-Penrose inverse using the rank equalities obtained in the preceding chapters. We also consider in the chapter how characterize some special types of matrices, such as, EP matrix, conjugate EP matrix, bi-EP matrix, star-dagger matrix, power-EP matrix, and so on.

In Chapter 7, we discuss various rank equalities for matrix expressions that involve two or more Moore-Penrose inverses, and then use them to characterize various matrix equalities that involve Moore-Penrose inverses.

In Chapter 8, we investigate various kind of reverse order laws for Moore-Penrose inverses of products of two or three matrices using the rank equalities established in the preceding chapters.

In Chapter 9, we investigate Moore-Penrose inverses of 2×2 block matrices, as well as $n\times n$ block matrices using the rank equalities established in the preceding chapters.

In Chapter 10, we investigate Moore-Penrose inverses of sums of matrices using the rank equalities established in the preceding chapters.

In Chapter 11, we study the relationships between Moore-Penrose inverses of block circulant matrices and sums of matrices. Based on them and the results in Chapter 9, we shall present a group of expressions for Moore-Penrose inverses of sums of matrices.

In Chapter 12, we present a group of formulas for expressing ranks of submatrices in the Moore-Penrose inverse of a matrix.

In Chapters 13—17, our work is concerned with rank equalities for Drazin inverses, group inverses, and weighted Moore-Penrose inverses of matrices and their applications. Various kinds of problems examined in Chapters 6—12 for Moore-Penrose inverses of matrices are almost considered in these five chapters for Drazin inverses, group inverses, and weighted Moore-Penrose inverses of matrices.

In Chapter 18, we present maximal and minimal ranks of the matrix expression $A - BXC$ with respect to the variant matrix X, and then consider rank and range invariance of $A - BXC$ with respect to the variant matrix X. In addition, we also consider shorted matrices of a matrix relative to a given matrix set.

In Chapter 19, we determine maximal and minimal ranks of the matrix expression $A - B_1X_1C_1 - B_2X_2C_2$ with respect to the variant matrices X_1 and X_2 under the conditions $R(B_1) \subseteq R(B_2)$ and $R(C_2^T) \subseteq R(C_1^T)$.

In Chapter 20, we determine maximal and minimal ranks of the matrix expression $A_1 - B_1XC_1$ subject to a consistent matrix equation $B_2XC_2 = A_2$, and consider some related topics.

In Chapters 21—23, we present maximal and minimal ranks of the Schur complement $D - CA^-B$ with respect to A^- and then consider various related topics, including some problems on generalized inverses of sums and products of matrices.

In Chapters 24—26, we determine extreme ranks of submatrices in solutions of the matrix equation $BXC = A$, and extreme ranks of two real matrices X_0 and X_1 in solutions to the complex matrix equation $B(X_0 + X_1)C = A$, as well as extreme ranks of solutions of the matrix equation $B_1XC_1 + B_2YC_2 = A$.

In Chapters 27—31, we consider extreme ranks of some general matrix expressions. The basic work is concerning extreme ranks of $A - B_1X_1C_1 - B_2X_2C_2$ with respect to the two independent variant matrices X_1 and X_2. The work is partially extended to some linear matrices expressions with more than two independent variant matrices. Quite a lot of consequences are derived from these results, including extreme ranks of $A_1 - B_1XC_1$ subject to a pair of consistent matrix equations $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$; extreme ranks of $A - BX - XC$ subject to a consistent matrix equations $BXC = D$; extreme ranks of a quadratic matrix expression $A - (A_1 - B_1X_1C_1)D(A_2 - B_2X_2C_2)$ with respect to X_1 and X_2. In addition, we also present many rank formulas for matrix expressions involving generalized inverses of matrices in these chapters.

Chapter 2

Basic rank formulas for Moore-Penrose inverses

The first and most fundamental rank formula used in the sequel is given below.

Theorem 2.1. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$, $C \in \mathcal{C}^{l\times n}$ and $D \in \mathcal{C}^{l\times k}$ be given. Then the rank of the Schur complement $S_A = D - CA^{\dagger}B$ satisfies the equality*

$$r(D - CA^{\dagger}B) = r\begin{bmatrix} A^*AA^* & A^*B \\ CA^* & D \end{bmatrix} - r(A). \tag{2.1}$$

Proof. It is obvious that

$$R(A^*B) \subseteq R(A^*) = R(A^*AA^*), \quad \text{and} \quad R(AC^*) \subseteq R(A) = R(AA^*A).$$

Then it follows by (1.7) and a well-known basic property $A^*(A^*AA^*)^{\dagger}A^* = A^{\dagger}$(see [118] pp. 69) that

$$r\begin{bmatrix} A^*AA^* & A^*B \\ CA^* & D \end{bmatrix} = r(A^*AA^*) + r[\, A - CA^*(A^*AA^*)^{\dagger}A^*B\,] = r(A) + r(D - CA^{\dagger}B),$$

establishing (2.1). $\square$

The significance of (2.1) is in that the rank of the Schur complement $S_A = D - CA^{\dagger}B$ can be evaluated by a block matrix formed by A, B, C and D in it, where no restrictions are imposed on S_A and no Moore-Penrose inverses appear in the right-hand side of (2.1). Thus (2.1) in fact provides us a powerful tool to express ranks of matrix expressions that involve Moore-Penrose inverses of matrices.

Eq. (2.1) can be extended to various general formulas. We next present some of them, which will widely be used in the sequel.

Theorem 2.2. *Let A_1, A_2, B_1, B_2, C_1, C_2 and D are matrices such that expression $D - C_1A_1^{\dagger}B_1 - C_2A_2^{\dagger}B_2$ is defined. Then*

$$r(D - C_1A_1^{\dagger}B_1 - C_2A_2^{\dagger}B_2) = r\begin{bmatrix} A_1^*A_1A_1^* & 0 & A_1^*B_1 \\ 0 & A_2^*A_2A_2^* & A_2^*B_2 \\ C_1A_1^* & C_2A_2^* & D \end{bmatrix} - r(A_1) - r(A_2). \tag{2.2}$$

Inparticular, if

$$R(B_1) \subseteq R(A_1), \quad R(C_1^*) \subseteq R(A_1^*), \quad R(B_2) \subseteq R(A_2) \quad \textit{and} \quad R(C_2^*) \subseteq R(A_2^*),$$

then

$$r(D - C_1A_1^{\dagger}B_1 - C_2A_2^{\dagger}B_2) = r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ C_1 & C_2 & D \end{bmatrix} - r(A_1) - r(A_2). \tag{2.3}$$

Let

$$C = [\, C_1, \ C_2\,], \qquad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}.$$

Then (2.1) can be written as (2.2), and (2.3) follows from (1.6). □

If the matrices in (2.2) satisfy certain conditions, the block matrix in (2.2) can easily be reduced to some simpler forms. Below are some of them.

Corollary 2.3. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$, $C \in \mathcal{C}^{l\times n}$ and $N \in \mathcal{C}^{k\times l}$ be given. Then*

$$r(N^{\dagger} - CA^{\dagger}B) = r\begin{bmatrix} AA^*A - A(BNC)^*A & AC^* & 0 \\ B^*A & 0 & N \\ 0 & N & 0 \end{bmatrix} - r(A) - r(N). \tag{2.4}$$

In particular, if $R(B^*A) \subseteq R(N)$ and $R(CA^*) \subseteq R(N^*)$, then

$$r(N^{\dagger} - CA^{\dagger}B) = r[\, AA^*A - A(BNC)^*A \,] + r(N) - r(A). \tag{2.5}$$

If $R(B^*A) \subseteq R(N)$, $R(CA^*) \subseteq R(N^*)$, $R(BN) \subseteq R(A)$ and $R[(NC)^*] \subseteq R(A^*)$, then

$$r(N^{\dagger} - CA^{\dagger}B) = r(A - BNC) + r(N) - r(A). \tag{2.6}$$

Theorem 2.4. *Let A_t, B_t, B_t, $C_t (t = 1,\, 2,\, \cdots,\, k)$ and D are matrices such that expression $D - C_1A_1^{\dagger}B_1 - \cdots - C_kA_k^{\dagger}B_k$ is defined. Then*

$$r(D - C_1A_1^{\dagger}B_1 - \cdots - C_kA_k^{\dagger}B_k) = r\begin{bmatrix} A^*AA^* & A^*B \\ CA^* & D \end{bmatrix} - r(A), \tag{2.7}$$

where $A = \mathrm{diag}(A_1,\, A_2,\, \cdots,\, A_k)$, $B^* = [\, B_1^*,\, B_2^*,\, \cdots,\, B_k^* \,]$ and $C = [\, C_1,\, C_2,\, \cdots,\, C_k \,]$.

Theorem 2.5. *Let A, B, C, D, P and Q are matrices such that expression $D - CP^{\dagger}AQ^{\dagger}B$ is defined. Then*

$$r(D - CP^{\dagger}AQ^{\dagger}B) = r\begin{bmatrix} P^*AQ^* & P^*PP^* & 0 \\ Q^*QQ^* & 0 & Q^*B \\ 0 & CP^* & -D \end{bmatrix} - r(P) - r(Q). \tag{2.8}$$

In particular, if

$$R(A) \subseteq R(P), \quad R(A^*) \subseteq R(Q^*), \quad R(B) \subseteq R(Q) \quad \text{and} \quad R(C^*) \subseteq R(P^*),$$

then

$$r(D - CP^{\dagger}AQ^{\dagger}B) = r\begin{bmatrix} A & P & 0 \\ Q & 0 & B \\ 0 & C & -D \end{bmatrix} - r(P) - r(Q). \tag{2.9}$$

Proof. Note that

$$\begin{aligned} r(D - CP^{\dagger}AQ^{\dagger}B) &= r\begin{bmatrix} A & AQ^{\dagger}B \\ CP^{\dagger}A & D \end{bmatrix} - r(A) \\ &= r\left(\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} + \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} 0 & P \\ Q & 0 \end{bmatrix}^{\dagger} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \right) - r(A). \end{aligned}$$

Applying (2.1) to it and then simplifying yields (2.8). Eq.(2.9) is derived from (2.8) by the rank cancellation law (1.8). □

Theorem 2.6. *Suppose that the matrix expression $S = D - C_1P_1^{\dagger}A_1Q_1^{\dagger}B_1 - C_2P_2^{\dagger}A_2Q_2^{\dagger}B_2$ is defined. Then*

$$r(S) = r\begin{bmatrix} P_1^*A_1Q_1^* & 0 & P_1^*P_1P_1^* & 0 & 0 \\ 0 & P_2^*A_2Q_2^* & 0 & P_2^*P_2P_2^* & 0 \\ Q_1^*Q_1Q_1^* & 0 & 0 & 0 & Q_1^*B_1 \\ 0 & Q_2^*Q_2Q_2^* & 0 & 0 & Q_2^*B_2 \\ 0 & 0 & C_1P_1^* & C_2P_2^* & -D \end{bmatrix} - d, \tag{2.10}$$

where $d = r(P_1) + r(P_2) + r(Q_1) + r(Q_2)$. *In particular, if*

$$R(A_i) \subseteq R(P_i), \quad R(A_i^*) \subseteq R(Q_i^*), \quad R(B_i) \subseteq R(Q_i) \;\; \text{and} \;\; R(C_i^*) \subseteq R(P_i^*), i = 1, \; 2,$$

then

$$r(S) = r\begin{bmatrix} A_1 & 0 & P_1 & 0 & 0 \\ 0 & A_2 & 0 & P_2 & 0 \\ Q_1 & 0 & 0 & 0 & B_1 \\ 0 & Q_2 & 0 & 0 & B_2 \\ 0 & 0 & C_1 & C_2 & -D \end{bmatrix} - r(P_1) - r(Q_1) - r(P_2) - r(Q_2). \tag{2.11}$$

Moreover,

$$r(\, D^\dagger - CP^\dagger AQ^\dagger B \,) = r\begin{bmatrix} D^*DD^* & 0 & 0 & D^* \\ 0 & P^*AQ^* & P^*PP^* & 0 \\ 0 & Q^*QQ^* & 0 & Q^*B \\ D^* & 0 & CP^* & 0 \end{bmatrix} - r(P) - r(Q) - r(D). \tag{2.12}$$

Proof. Writing S as

$$S = D - [\, C_1, \; C_2 \,] \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}^\dagger \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix}^\dagger \begin{bmatrix} B_1 \\ B_2 \end{bmatrix},$$

and then applying (2.8) to it produce (2.10). Eq. (2.11) is derived from (2.10) by the rank cancellation law (1.8). Eq. (2.12) is a special case of (2.10). $\square$

It is easy to see that a general rank formula for

$$D - C_1P_1^\dagger A_1Q_1^\dagger B_1 - C_2P_2^\dagger A_2Q_2^\dagger B_2 - \cdots - C_kP_2^\dagger A_kQ_k^\dagger B_k$$

can also be established by the similar method for deriving (2.10). As to some other general matrix expressions, such as

$$S_k = A_0P_1^\dagger A_1P_2^\dagger A_2 \cdots P_k^\dagger A_k$$

and their linear combinations, the formulas for expressing their ranks can also be established. However they are quite tedious in form, we do not intend to give them here.

Chapter 3

Rank equalities for idempotent matrices

A square matrix A is said to be idempotent if $A^2 = A$. If we consider it as a matrix equation $A^2 = A$, then its general solution can be written as $A = V(V^2)^{\dagger}V$, where V is an arbitrary square complex matrix. This assertion can easily be verified. In fact, $A = V(V^2)^{\dagger}V$, apparently satisfies $A^2 = A$. Now for any matrix A with $A^2 = A$, we let $V = A$. Then $V(V^2)^{\dagger}V = A(A^2)^{\dagger}A = AA^{\dagger}A = A$. Thus $A = V(V^2)^{\dagger}V$ is indeed the general solution the idempotent equation $A^2 = A$. This fact clearly implies that any matrix expression that involves idempotent matrices could be regarded as a conventional matrix expression that involves Moore-Penrose inverses of matrices. Thus the formulas in Chapter 2 are all applicable to determine ranks of matrix expressions that involve idempotent matrices. However because of speciality of idempotent matrices, the rank equalities related to idempotent matrices can also be deduced by various elementary methods. The results in the chapter are originally derived by the rank formulas in Chapter 2, we later also find some elementary methods to establish them. So we only show these results in these elementary methods inn this chapter.

Theorem 3.1. *Let $P, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the difference $P - Q$ satisfies the rank equalities*

$$r(P - Q) = r\begin{bmatrix} P \\ Q \end{bmatrix} + r[\,P,\ Q\,] - r(P) - r(Q), \tag{3.1}$$

$$r(P - Q) = r(P - PQ) + r(PQ - Q), \tag{3.2}$$

$$r(P - Q) = r(P - QP) + r(QP - Q). \tag{3.3}$$

Proof. Let $M = \begin{bmatrix} -P & 0 & P \\ 0 & Q & Q \\ P & Q & 0 \end{bmatrix}$. Then it is easy to see by block elementary operations of matrices that

$$r(M) = r\begin{bmatrix} -P & 0 & 0 \\ 0 & Q & 0 \\ 0 & 0 & P - Q \end{bmatrix} = r(P) + r(Q) + r(P - Q).$$

On the other hand, note that $P^2 = P$ and $Q^2 = Q$. It is also easy to find by block elementary operations of matrices that

$$r(M) = r\begin{bmatrix} -P & 0 & P \\ -QP & 0 & Q \\ P & Q & 0 \end{bmatrix} = r\begin{bmatrix} 0 & 0 & P \\ 0 & 0 & Q \\ P & Q & 0 \end{bmatrix} = r\begin{bmatrix} P \\ Q \end{bmatrix} + r[\,P,\ Q\,].$$

Combining the above two equalities yields (3.1). Consequently applying (1.2) and (1.3) to $[\,P,\ Q\,]$ and $\begin{bmatrix} P \\ Q \end{bmatrix}$ in (3.1) respectively yields

$$r[\,P,\ Q\,] = r(P) + r(Q - PQ), \tag{3.4}$$

$$r[\,P,\ Q\,] = r(Q) + r(\,P - QP\,), \tag{3.5}$$

$$r\begin{bmatrix} P \\ Q \end{bmatrix} = r(P) + r(\,Q - QP\,), \tag{3.6}$$

$$r\begin{bmatrix} P \\ Q \end{bmatrix} = r(Q) + r(\,P - PQ\,). \tag{3.7}$$

Putting (3.4) and (3.7) in (3.1) produces (3.2), putting (3.5) and (3.6) in (3.1) produces (3.3). □

Corollary 3.2. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then*

(a) $R(\,P - PQ\,) \cap R(\,PQ - Q\,) = \{0\}$ *and* $R[(\,P - PQ\,)^*] \cap R[(\,PQ - Q\,)^*] = \{0\}$.

(b) $R(\,P - QP\,) \cap R(\,QP - Q\,) = \{0\}$ *and* $R[(\,P - QP\,)^*] \cap R[(\,QP - Q\,)^*] = \{0\}$.

(c) *If* $PQ = 0$ *or* $QP = 0$, *then* $r(\,P - Q\,) = r(P) + r(Q)$, *i.e.*, $R(P) \cap R(Q) = \{0\}$ *and* $R(P^*) \cap R(Q^*) = \{0\}$.

(d) *If both P and Q are Hermitian idempotent, then* $r(\,P - Q\,) = 2r[\,P,\ Q\,] - r(P) - r(Q)$.

Proof. Parts (a) and (b) follows from applying Lemma 1.4(d) to (3.2) and (3.3). Part (c) is a direct consequence of (3.2) and (3.3). Part (d) follows from (3.1). □

On the basis of (3.1), we can easily deduce the following known result due to Hartwig and Styan [66] on the rank subtractivity two idempotent matrices.

Corollary 3.3. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the following statements are equivalent*:

(a) $r(\,P - Q\,) = r(P) - r(Q)$, *i.e.*, $Q \leq_{rs} P$.

(b) $r\begin{bmatrix} P \\ Q \end{bmatrix} = r[\,P,\ Q\,] = r(P)$.

(c) $R(Q) \subseteq R(P)$ *and* $R(Q^*) \subseteq R(P^*)$.

(d) $PQ = QP = Q$.

(e) $PQP = Q$.

Proof. The equivalence of Parts (a) and (b) follows immediately from applying (3.1). The equivalence of Parts (b), (c), (d) and (e) can trivially be verified by (1.2) and (1.3). □

Corollary 3.4. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the following statements are equivalent*:

(a) *The difference $P - Q$ is nonsingular.*

(b) $r\begin{bmatrix} P \\ Q \end{bmatrix} = r[\,P,\ Q\,] = r(P) + r(Q) = m$.

(c) $R(P) \oplus R(Q) = R(P^*) \oplus R(Q^*) = \mathcal{C}^m$.

Proof. Follows directly from (3.1). □

Notice that if a matrix P is idempotent, the $I_m - P$ is also idempotent. Thus replacing P in (3.1) by $I_m - P$, we get the following.

Theorem 3.5. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the rank of $I_m - P - Q$ satisfies the equalities*

$$r(\,I_m - P - Q\,) = r(PQ) + r(QP) - r(P) - r(Q) + m, \tag{3.8}$$
$$r(\,I_m - P - Q\,) = r(\,I_m - P - Q + PQ\,) + r(PQ), \tag{3.9}$$
$$r(\,I_m - P - Q\,) = r(\,I_m - P - Q + QP\,) + r(QP). \tag{3.10}$$

Proof. Replacing P in (3.1) by $I_m - P$ yields

$$r(\,I_m - P - Q\,) = r\begin{bmatrix} I_m - P \\ Q \end{bmatrix} + r[\,I_m - P,\ Q\,] - r(\,I_m - P\,) - r(Q). \tag{3.11}$$

It follows by (1.2) and (1.3) that

$$r[\, I_m - P,\ Q\,] = r(\, I_m - P\,) + r[\, Q - (\, I_m - P\,)Q\,] = m - r(P) + r(PQ),$$

and

$$r\begin{bmatrix} I_m - P \\ Q \end{bmatrix} = r(\, I_m - P\,) + r[\, Q - Q(\, I_m - P\,)\,] = m - r(P) + r(QP).$$

Putting them in (3.11) produces (3.8). On the other hand, replacing P in (3.2) and (3.3) by $I_m - P$ produces

$$\begin{aligned} r[\,(I_m - P) - Q\,] &= r[\,(I_m - P) - (I_m - P)Q\,] + r[\,(I_m - P)Q - Q\,] \\ &= r(\, I_m - P - Q + PQ\,) + r(PQ), \end{aligned}$$

and

$$\begin{aligned} r[\,(I_m - P) - Q\,] &= r[\,(I_m - P) - Q(I_m - P)\,] + r[\,Q(I_m - P) - Q\,] \\ &= r(\, I_m - P - Q + QP\,) + r(QP), \end{aligned}$$

both of which are exactly (3.9) and (3.10). □

Corollary 3.6. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then*

(a) $R(\, I_m - P - Q + PQ\,) \cap R(PQ) = \{0\}$ *and* $R[(\, I_m - P - Q + PQ\,)^*] \cap R[(PQ)^*] = \{0\}$.

(b) $R(\, I_m - P - Q + QP\,) \cap R(QP) = \{0\}$ *and* $R[(\, I_m - P - Q + QP\,)^*] \cap R[(QP)^*] = \{0\}$.

(c) $P + Q = I_m \ \Leftrightarrow\ PQ = QP = 0$ *and* $R(P) \oplus R(Q) = R(P^*) \oplus R(Q^*) = \mathcal{C}^m$.

(d) *If* $PQ = QP = 0$, *then* $r(\, I_m - P - Q\,) = m - r(P) - r(Q)$.

(e) $I_m - P - Q$ *is nonsingular if and only if* $r(PQ) = r(QP) = r(P) = r(Q)$.

(f) *If both P and Q are Hermitian idempotent, then* $r(\, I_m - P - Q\,) = 2r(PQ) - r(P) - r(Q) + m$.

Proof. Parts (a) and (b) follow from applying Lemma 1.4(d) to (3.9) and (3.10). Note from (3.8)—(3.10) that $P + Q = I_m$ is equivalent to $PQ = QP = 0$ and $r(P) + r(Q) = m$. This assertion is also equivalent to $PQ = QP = 0$ and $R(P) \oplus R(Q) = R(P^*) \oplus R(Q^*) = \mathcal{C}^m$, which is Part (c). Parts (d), (e) and (f) follow from (3.8). □

As for the rank of sum of two idempotent matrices, we have the following several results.

Theorem 3.7. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the sum $P + Q$ satisfies the rank equalities*

$$r(\, P + Q\,) = r\begin{bmatrix} P & Q \\ Q & 0 \end{bmatrix} - r(Q) = r\begin{bmatrix} Q & P \\ P & 0 \end{bmatrix} - r(P), \tag{3.12}$$

$$r(\, P + Q\,) = r(\, P - PQ - QP + QPQ\,) + r(Q), \tag{3.13}$$

$$r(\, P + Q\,) = r(\, Q - PQ - QP + PQP\,) + r(P). \tag{3.14}$$

Proof. Let $M = \begin{bmatrix} P & 0 & P \\ 0 & Q & Q \\ P & Q & 0 \end{bmatrix}$. Then it is easy to see by block elementary operations of matrices that

$$r(M) = r\begin{bmatrix} P & 0 & 0 \\ 0 & Q & 0 \\ 0 & 0 & -P - Q \end{bmatrix} = r(P) + r(Q) + r(\, P + Q\,).$$

On the other hand, note that $P^2 = P$ and $Q^2 = Q$. It is also easy to find by block elementary operations of matrices that

$$r(M) = r\begin{bmatrix} P & 0 & P \\ -QP & 0 & Q \\ P & Q & 0 \end{bmatrix} = r\begin{bmatrix} 2P & 0 & P \\ 0 & 0 & Q \\ P & Q & 0 \end{bmatrix} = r\begin{bmatrix} 2P & 0 & 0 \\ 0 & 0 & Q \\ 0 & Q & \frac{1}{2}P \end{bmatrix} = r\begin{bmatrix} P & Q \\ Q & 0 \end{bmatrix} + r(P),$$

and

$$r(M) = r\begin{bmatrix} 0 & -PQ & P \\ 0 & Q & Q \\ P & Q & 0 \end{bmatrix} = r\begin{bmatrix} 0 & 0 & P \\ 0 & 2Q & Q \\ P & Q & 0 \end{bmatrix} = r\begin{bmatrix} 0 & 0 & P \\ 0 & 2Q & 0 \\ P & 0 & \frac{1}{2}Q \end{bmatrix} = r\begin{bmatrix} Q & P \\ P & 0 \end{bmatrix} + r(Q).$$

The combination of the above three rank equalities yields the two equalities in (3.12). Consequently applying (1.4) to the two block matrices in (3.12) yields (3.13) and (3.14), respectively. □

Corollary 3.8. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices.*

(a) *If $PQ = QP$, then*

$$r(\,P + Q\,) = r[\,P,\ Q\,] = r\begin{bmatrix} P \\ Q \end{bmatrix}, \tag{3.15}$$

orequivalently,

$$R(Q) \subseteq R(P + Q) \quad and \quad R(Q^*) \subseteq R(P^* + Q^*). \tag{3.16}$$

(b) *If $R(Q) \subseteq R(P)$ or $R(Q^*) \subseteq R(P^*)$, then $r(\,P + Q\,) = r(P)$.*

Proof. If $PQ = QP$, then (3.13) and (3.14) reduce to

$$r(\,P + Q\,) = r(\,P - PQ\,) + r(Q) = r(\,Q - PQ\,) + r(P).$$

Combining them with (3.4) and (3.7) yields (3.15). The equivalence of (3.15) and (3.16) follows from a simple fact that

$$r\begin{bmatrix} P \\ Q \end{bmatrix} = r\begin{bmatrix} P+Q \\ Q \end{bmatrix} \quad \text{and} \quad r[\,P,\ Q\,] = r[\,P + Q,\ Q\,],$$

as well as Lemma 1.2(c) and (d). The result in Part (b) follows immediately from (3.12). □

Corollary 3.9. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the following five statements are equivalent*:

(a) *The sum $P + Q$ is nonsingular.*

(b) $r\begin{bmatrix} P \\ Q \end{bmatrix} = m$ *and* $R\begin{bmatrix} P \\ Q \end{bmatrix} \cap R\begin{bmatrix} Q \\ 0 \end{bmatrix} = \{0\}$.

(c) $r[\,P,\ Q\,] = m$ *and* $R\begin{bmatrix} P^* \\ Q^* \end{bmatrix} \cap R\begin{bmatrix} Q^* \\ 0 \end{bmatrix} = \{0\}$.

(d) $r\begin{bmatrix} Q \\ P \end{bmatrix} = m$ *and* $R\begin{bmatrix} Q \\ P \end{bmatrix} \cap R\begin{bmatrix} P \\ 0 \end{bmatrix} = \{0\}$.

(e) $r[\,Q,\ P\,] = m$ *and* $R\begin{bmatrix} Q^* \\ P^* \end{bmatrix} \cap R\begin{bmatrix} P^* \\ 0 \end{bmatrix} = \{0\}$.

Proof. In light of (3.12), the sum $P + Q$ is nonsingular if and only if

$$r\begin{bmatrix} P & Q \\ Q & 0 \end{bmatrix} = r(Q) + m, \tag{3.17}$$

or equivalently

$$r\begin{bmatrix} Q & P \\ P & 0 \end{bmatrix} = r(P) + m. \tag{3.18}$$

Observe that

$$r\begin{bmatrix} P & Q \\ Q & 0 \end{bmatrix} \leq r\begin{bmatrix} P \\ Q \end{bmatrix} + r\begin{bmatrix} Q \\ 0 \end{bmatrix} \leq m + r(Q),$$

$$r\begin{bmatrix} P & Q \\ Q & 0 \end{bmatrix} \leq r[\,P,\ Q\,] + r[\,Q,\ 0\,] \leq m + r(Q),$$

$$r\begin{bmatrix} Q & P \\ P & 0 \end{bmatrix} \leq r\begin{bmatrix} Q \\ P \end{bmatrix} + r\begin{bmatrix} P \\ 0 \end{bmatrix} \leq m + r(P),$$

$$r\begin{bmatrix} Q & P \\ P & 0 \end{bmatrix} \leq r[\,Q,\ P\,] + r[\,P,\ 0\,] \leq m + r(P).$$

Combining them with (3.17) and (3.18) yields the equivalence of Parts (a)—(e). □

Theorem 3.10. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then*

(a) *The rank of* $I_m + P - Q$ *satisfies the equality*

$$r(\, I_m + P - Q\,) = r(QPQ) - r(Q) + m. \tag{3.19}$$

(b) *The rank of* $2I_m - P - Q$ *satisfies the two equalities*

$$r(\, 2I_m - P - Q\,) = r(\, Q - QPQ\,) - r(Q) + m, \tag{3.20}$$
$$r(\, 2I_m - P - Q\,) = r(\, P - PQP\,) - r(P) + m. \tag{3.21}$$

Proof. Replacing Q in (3.12) by the idempotent matrix $I_m - Q$ and applying (1.4) to it yields

$$\begin{aligned} r(\, I_m + P - Q\,) &= r\begin{bmatrix} P & I_m - Q \\ I_m - Q & 0 \end{bmatrix} - r(\, I_m - Q\,) \\ &= r(\, I_m - Q\,) + r[(\, I_m - (\, I_m - Q\,))P(\, I_m - (\, I_m - Q\,))] \\ &= m - r(Q) + r(QPQ), \end{aligned}$$

establishing (3.19). Further, replacing P and Q in (3.12) by $I_m - P$ and $I_m - Q$, we also by (1.4) find that

$$\begin{aligned} r(\, 2I_m - P - Q\,) &= r\begin{bmatrix} I_m - P & I_m - Q \\ I_m - Q & 0 \end{bmatrix} - r(\, I_m - Q\,) \\ &= r(\, I_m - Q\,) + r[\,(\, I_m - (\, I_m - Q\,))(\, I_m - P\,)(\, I_m - (\, I_m - Q\,))\,] \\ &= m - r(Q) + r(\, Q - QPQ\,), \end{aligned}$$

establishing (3.20). Similarly, we can show (3.21). $\square$

Corollary 3.11. *Let* $P,\, Q \in \mathcal{C}^{m\times m}$ *be two idempotent matrices.*

(a) *If* $R(P) \subseteq R(Q)$ *and* $R(P^*) \subseteq R(Q^*)$, *then* P *and* Q *satisfy the two rank equalities*

$$r(\, I_m + P - Q\,) = m + r(P) - r(Q), \tag{3.22}$$
$$r(\, 2I_m - P - Q\,) = m + r(\, Q - P\,) - r(Q). \tag{3.23}$$

(b) $I_m + P - Q$ *is nonsingular* $\Leftrightarrow r(QPQ) = r(Q)$.

(c) $2I_m - P - Q$ *is nonsingular* $\Leftrightarrow r(\, P - PQP\,) = r(P) \Leftrightarrow r(\, Q - QPQ\,) = r(Q)$.

(d) $Q - P = I_m \Leftrightarrow r(QPQ) + r(Q) = m$.

Proof. The two conditions $R(P) \subseteq R(Q)$ and $R(P^*) \subseteq R(Q^*)$ are equivalent to $QP = P = PQ$. In that case, (3.19) reduces to (3.22), (3.20) and (3.21) reduce to (3.23). The results in Parts (a)—(c) are direct consequences of (3.19). $\square$

we next consider the rank of $PQ - QP$ for two idempotent matrices P and Q.

Theorem 3.12. *Let* $P,\, Q \in \mathcal{C}^{m\times m}$ *be two idempotent matrices. Then the difference* $PQ - QP$ *satisfies the five rank equalities*

$$r(\, PQ - QP\,) = r(\, P - Q\,) + r(\, I_m - P - Q\,) - m, \tag{3.24}$$
$$r(\, PQ - QP\,) = r(\, P - Q\,) + r(PQ) + r(QP) - r(P) - r(Q), \tag{3.25}$$
$$r(\, PQ - QP\,) = r\begin{bmatrix} P \\ Q \end{bmatrix} + r[\, P,\ Q\,] + r(PQ) + r(QP) - 2r(P) - 2r(Q), \tag{3.26}$$
$$r(\, PQ - QP\,) = r(\, P - PQ\,) + r(\, PQ - Q\,) + r(PQ) + r(QP) - r(P) - r(Q), \tag{3.27}$$
$$r(\, PQ - QP\,) = r(\, P - QP\,) + r(\, QP - Q\,) + r(PQ) + r(QP) - r(P) - r(Q). \tag{3.28}$$

In particular, if both P *and* Q *are Hermitian idempotent, then*

$$r(\, PQ - QP\,) = 2r[\, P,\ Q\,] + 2r(PQ) - 2r(P) - 2r(Q). \tag{3.29}$$

Proof. It is easy to verify that that $PQ - QP = (\, P - Q\,)(\, P + Q - I_m\,)$. Thus the rank of $PQ - QP$ can be expressed as

$$r(\, PQ - QP\,) = r[\,(\, P - Q\,)(\, P + Q - I_m\,)\,] = r\begin{bmatrix} I_m & P + Q - I_m \\ P - Q & 0 \end{bmatrix} - m. \tag{3.30}$$

On the other hand, it is easy to verify the factorization

$$\begin{bmatrix} I_m & P+Q-I_m \\ P-Q & 0 \end{bmatrix} = \begin{bmatrix} I_m & 2P-I_m \\ 0 & I_m \end{bmatrix} \begin{bmatrix} 0 & P+Q-I_m \\ P-Q & 0 \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 2Q-I_m & I_m \end{bmatrix}.$$

Hence

$$r\begin{bmatrix} I_m & P+Q-I_m \\ P-Q & 0 \end{bmatrix} = r(\, P-Q\,) + r(\, I_m - P - Q\,).$$

Putting it in (3.30) yields (3.24). Consequently putting (3.8) in (3.24) yields (3.25); putting (3.1) in (3.25) yields (3.26); putting (3.2) and (3.3) respectively in (3.25) yields (3.27) and (3.28). □

Corollary 3.13. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the following five statements are equivalent*:

(a) $PQ = QP$.
(b) $r(\, P-Q\,) + r(\, I_m - P - Q\,) = m$.
(c) $r(\, P-Q\,) = r(P) + r(Q) - r(PQ) - r(QP)$.
(d) $r(\, P - PQ\,) = r(P) - r(PQ)$ *and* $r(\, Q - PQ\,) = r(Q) - r(PQ)$, *i.e.*, $PQ \leq_{rs} P$ *and* $PQ \leq_{rs} Q$.
(e) $r(\, P - QP\,) = r(P) - r(QP)$ *and* $r(\, Q - QP\,) = r(Q) - r(QP)$, *i.e.*, $QP \leq_{rs} P$ *and* $QP \leq_{rs} Q$.
(f) $r\begin{bmatrix} P \\ Q \end{bmatrix} = r(P) + r(Q) - r(PQ)$ *and* $r[\, P,\ Q\,] = r(P) + r(Q) - r(QP)$.
(g) $r\begin{bmatrix} P \\ Q \end{bmatrix} = r(P) + r(Q) - r(QP)$ *and* $r[\, P,\ Q\,] = r(P) + r(Q) - r(PQ)$.

Proof. Follows immediately from (3.24)—(3.28). □

Corollary 3.14. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the following three statements are equivalent:*

(a) $r(\, PQ - QP\,) = r(\, P - Q\,)$.
(b) $I_m - P - Q$ *is nonsingular.*
(c) $r(PQ) = r(QP) = r(P) = r(Q)$.

Proof. The equivalence of Parts (a) and (b) follows from (3.24). The equivalence of Parts (b) and (c) follows from Corollary 3.6(e). □

Corollary 3.15. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the following three statements are equivalent:*

(a) $PQ - QP$ *is nonsingular.*
(b) $P - Q$ *and* $I_m - P - Q$ *are nonsingular.*
(c) $R(P) \oplus R(Q) = R(P^*) \oplus R(Q^*) = \mathcal{C}^m$ *and* $r(PQ) = r(QP) = r(P) = r(Q)$ *hold.*

Proof. The equivalence of Parts (a) and (b) follows from (3.24). The equivalence of Parts (b) and (c) follows from Corollaries 3.4(e) and 3.6(e). □

A group of analogous rank equalities can also be derived for $PQ + QP$, where P and Q are two idempotent matrices P and Q.

Theorem 3.16. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then $PQ{+}QP$ satisfies the rank equalities*

$$r(\, PQ + QP\,) = r(\, P + Q\,) + r(\, I_m - P - Q\,) - m, \tag{3.31}$$
$$r(\, PQ + QP\,) = r(\, P + Q\,) + r(PQ) + r(QP) - r(P) - r(Q), \tag{3.32}$$
$$r(\, PQ + QP\,) = r(\, P - PQ - QP + QPQ\,) + r(PQ) + r(QP) - r(P), \tag{3.33}$$
$$r(\, PQ + QP\,) = r(\, Q - PQ - QP + PQP\,) + r(PQ) + r(QP) - r(Q). \tag{3.34}$$

Proof. Note that $PQ + QP = (\, P + Q\,)^2 - (\, P + Q\,)$. Then applying (1.11) to it, we directly obtain (3.31). Consequently, putting (3.8) in (3.21) yields (3.32), putting (3.13) and (3.14) respectively in (3.32) yields (3.33) and (3.34). □

Corollary 3.17. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the following four statements are equivalent*:

(a) $r(\, PQ + QP\,) = r(\, P + Q\,)$.
(b) $I_m - P - Q$ *is nonsingular.*
(c) $r(PQ) = r(QP) = r(P) = r(Q)$.
(d) $r(\, PQ - QP\,) = r(\, P - Q\,)$.

Proof. The equivalence of Parts (a) and (b) follows from (3.31), and the equivalence of Parts (b)—(d) comes from Corollary 3.14. □

Corollary 3.18. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the following two statements are equivalent*:
(a) $PQ + QP$ *is nonsingular.*
(b) $P + Q$ *and* $I_m - P - Q$ *are nonsingular.*

Proof. Follows directly from (3.31). □

Combining the two rank equalities in (3.24) and (3.31), we obtain the following.

Corollary 3.19. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then both of them satisfy the following rank identity*

$$r(\, P + Q\,) + r(\, PQ - QP\,) = r(\, P - Q\,) + r(\, PQ + QP\,). \tag{3.35}$$

Thoerem 3.20. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then*

$$r[\,(\, P - Q\,)^2 - (\, P - Q\,)\,] = r(\, I_m - P + Q\,) + r(\, P - Q\,) - m. \tag{3.36}$$

$$r[\,(\, P - Q\,)^2 - (\, P - Q\,)\,] = r(PQP) - r(P) + r(\, P - Q\,). \tag{3.37}$$

Proof. Eq. (3.36) is derived from (1.11). According to (3.19), we have $r(\, I_m - P + Q\,) = r(PQP) - r(P) + m$. Putting it in (3.36) yields (3.37). □

Corollary 3.21 (Hartwig and Styan [66]). *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the following five statements are equivalent*:
(a) $P - Q$ *is idempotent.*
(b) $r(\, I_m - P + Q\,) = m - r(\, P - Q\,)$.
(c) $r(\, P - Q\,) = r(P) - r(Q)$, *i.e.*, $Q \leq_{rs} P$.
(d) $R(Q) \subseteq R(P)$ *and* $R(Q^*) \subseteq R(P^*)$.
(e) $PQP = Q$.

Proof. The equivalence of Parts (a) and (b) follows immediately from (3.36), and the equivalence of Parts (c), (d) and (e) is from Corollary 3.3(d). The equivalence of Parts (a) and (e) follows from a direct matrix computation. □

In Chapter 4, we shall also establish a rank formula for $(\, P - Q\,)^3 - (\, P - Q\,)$ and consider tripotency of $P - Q$, where $P,\ Q$ are two idempotent matrices.

Theorem 3.22. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then $I_m - PQ$ satisfies the rank equalities*

$$r(\, I_m - PQ\,) = r(\, 2I_m - P - Q\,) = r[\,(\, I_m - P\,) + (\, I_m - Q\,)\,]. \tag{3.38}$$

Proof. According to (1.10) we have

$$r(\, I_m - PQ\,) = r(\, Q - QPQ\,) - r(Q) + m.$$

Consequently putting (3.20) in it yields (3.38). □

Corollary 3.23. *Let $P,\, Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then the sum $P + Q$ satisfies the rank identities*

$$r(\, P + Q\,) = r(\, P + Q - PQ\,) = r(\, P + Q - QP\,). \tag{3.39}$$

In particular, if $PQ = QP$, then

$$r(\, P + Q\,) = r(P) + r(Q) - r(PQ). \tag{3.40}$$

Proof. Replacing P and Q in (3.38) by two idempotent matrices $I_m - P$ and $I_m - Q$ immediately yields (3.39). If $PQ = QP$, then we know by (3.13) and (3.14) that

$$r(P + Q) = r(P - PQ) + r(Q) = r(Q - QP) + r(P), \tag{3.41}$$

and by Corollary 3.13 we also know that

$$r(P - PQ) = r(P) - r(PQ) \quad \text{and} \quad r(Q - QP) = r(Q) - r(QP). \tag{3.42}$$

Putting (3.42) in (3.41) yields (3.40). □

Corollary 3.24. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then*

$$r[PQ - (PQ)^2] = r(I_m - PQ) + r(PQ) - m = r(2I_m - P - Q) + r(PQ) - m. \tag{3.43}$$

In particular, the following statements are equivalent:
- (a) PQ *is idempotent.*
- (b) $r(I_m - PQ) = m - r(PQ)$.
- (c) $r(2I_m - P - Q) = m - r(PQ)$.

Proof. Applying (1.11) to $PQ - (PQ)^2$ gives the first equality in (3.43). The second one follows from (3.38). □

Corollary 3.25. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then*

$$r(I_m - P - Q + PQ) = m - r(P) - r(Q) + r(QP).$$

Proof. This follows from replacing A in (1.4) by I_m. □

Notice that if a matrix A is idempotent, then A^* is also idempotent. Thus we can easily find the following.

Corollary 3.26. *Let $P \in \mathcal{C}^{m\times m}$ be an idempotent matrix. Then*
- (a) $r(P - P^*) = 2r[P,\ P^*] - 2r(P)$.
- (b) $r(I_m - P - P^*) = r(I_m + P - P^*) = m$.
- (c) $r(P + P^*) = r(PP^* + P^*P) = r[P,\ P^*]$, *i.e.*, $R(P) \subseteq R(P + P^*)$ *and* $R(P^*) \subseteq R(P + P^*)$.
- (d) $r(PP^* - P^*P) = r(P - P^*)$.

Proof. Part (a) follows from (3.1). Part (b) follows from (3.8) and (3.22). Part (c) follows from (3.31). Part (d) follows from (3.24) and Part (b). □

The results in the preceding theorems and corollaries can easily be extended to matrices with properties $P^2 = \lambda P$ and $Q^2 = \mu Q$, where $\lambda \neq 0$ and $\mu \neq 0$. In fact, observe that

$$\left(\frac{1}{\lambda} P \right)^2 = \frac{1}{\lambda^2} P^2 = \frac{1}{\lambda} P, \qquad \left(\frac{1}{\mu} Q \right)^2 = \frac{1}{\mu^2} Q^2 = \frac{1}{\mu} Q.$$

Thus both P/λ and Q/μ are idempotent. In that case, applying the results in the previous theorems and corollaries, one may establish a variety of rank equalities and their consequences related to such kind of matrices. For example,

$$\begin{aligned}
r(\mu P - \lambda Q) &= r \begin{bmatrix} P \\ Q \end{bmatrix} + r[P,\ Q] - r(P) - r(Q), \\
r(\mu P + \lambda Q) &= r \begin{bmatrix} P & Q \\ Q & 0 \end{bmatrix} - r(Q) = r \begin{bmatrix} Q & P \\ P & 0 \end{bmatrix} - r(P), \\
r(\lambda\mu I_m - \mu P - \lambda Q) &= r(PQ) + r(QP) - r(P) - r(Q) + m, \\
r(PQ - QP) &= r(\mu P - \lambda Q) + r(\lambda\mu I_m - \mu P - \lambda Q) - m, \\
r(PQ + QP) &= r(\mu P + \lambda Q) + r(\lambda\mu I_m - \mu P - \lambda Q) - m, \\
r(\lambda\mu I_m - PQ) &= r(2\lambda\mu I_m - \mu P - \lambda Q),
\end{aligned}$$

and so on. We do not intend to present them in details.

Chapter 4

More on rank equalities for idempotent matrices

The rank equalities in Chapter 3 can partially be extended to matrix expressions that involve idempotent matrices and general matrices. In addition, they can also be applied to establish rank equalities related to involutory matrices. The corresponding results are presented in this chapter.

Theorem 4.1. *Let $A \in \mathcal{C}^{m\times n}$ be given, $P \in \mathcal{C}^{m\times m}$ and $Q \in \mathcal{C}^{n\times n}$ be two idempotent matrices. Then the difference $PA - AQ$ satisfies the two rank equalities*

$$r(\,PA - AQ\,) = r\begin{bmatrix} PA \\ Q \end{bmatrix} + r[\,AQ, \ P\,] - r(P) - r(Q), \tag{4.1}$$

$$r(\,PA - AQ\,) = r(\,PA - PAQ\,) + r(\,PAQ - AQ\,). \tag{4.2}$$

Proof. Let $M = \begin{bmatrix} -P & 0 & PA \\ 0 & Q & Q \\ P & AQ & 0 \end{bmatrix}$. Then it is easy to see by the block elementary operations of matrices that

$$r(M) = r\begin{bmatrix} -P & 0 & 0 \\ 0 & Q & 0 \\ 0 & 0 & PA - AQ \end{bmatrix} = r(P) + r(Q) + r(\,PA - AQ\,). \tag{4.3}$$

On the other hand, note that $P^2 = P$ and $Q^2 = Q$. It is also easy to find by block elementary operations of matrices that

$$r(M) = r\begin{bmatrix} 0 & PAQ & PA \\ 0 & Q & Q \\ P & AQ & 0 \end{bmatrix} = r\begin{bmatrix} 0 & 0 & PA \\ 0 & 0 & Q \\ P & AQ & 0 \end{bmatrix} = r\begin{bmatrix} PA \\ Q \end{bmatrix} + r[\,AQ, \ P\,]. \tag{4.4}$$

Combining (4.3) and (4.4) yields (4.1). Consequently applying (1.2) and (1.3) to $[\,AQ, \ P\,]$ and $\begin{bmatrix} PA \\ Q \end{bmatrix}$ in (4.1) respectively yields (4.2). $\square$

Corollary 4.2. *Let $A \in \mathcal{C}^{m\times n}$ be given, $P \in \mathcal{C}^{m\times m}$ and $Q \in \mathcal{C}^{n\times n}$ be two idempotent matrices. Then*

(a) $R(\,PA - PAQ\,) \cap R(\,PAQ - AQ\,) = \{0\}$ *and* $R[(\,PA - PAQ\,)^*] \cap R[(\,PAQ - AQ\,)^*] = \{0\}$.

(b) *If* $PAQ = 0$, *then* $r(\,PA - AQ\,) = r(PA) + r(AQ)$, *or, equivalently* $R(PA) \cap R(AQ) = \{0\}$ *and* $R[(PA)^*] \cap R[(AQ)^*] = \{0\}$.

(c) $PA = AQ \Leftrightarrow PA(\,I - Q\,) = 0$ *and* $(\,I - P\,)AQ = 0 \Leftrightarrow R(AQ) \subseteq R(P)$ *and* $R[(PA)^*] \subseteq R(Q^*)$.

Proof. Part (a) follows from applying Lemma 1.4(d) to (4.2). Parts (b) and (c) are direct consequences of (4.2). $\square$

Corollary 4.3. *Let $A, P, Q \in \mathcal{C}^{m\times m}$ be given with $P, \ Q$ being two idempotent matrices. Then the following three statements are equivalent:*

(a) $PA - AQ$ *is nonsingular.*

(b) $r\begin{bmatrix} PA \\ Q \end{bmatrix} = r[\,AQ,\ P\,] = r(P) + r(Q) = m.$

(c) $r(PA) = r(P),\ r(AQ) = r(Q)$ *and* $R(AQ) \oplus R(P) = R[(PA)^*] \oplus R(Q^*) = \mathcal{C}^m.$

Proof. Follows from (4.1). $\square$

Based on Corollary 4.2(c), we find an interesting result on the general solution of a matrix equation.

Corollary 4.4. *Let $P \in \mathcal{C}^{m\times m}$ and $Q \in \mathcal{C}^{n\times n}$ be two idempotent matrices. Then the general solution of the matrix equation $PX = XQ$ can be written in the two forms*

$$X = PUQ + (\,I_m - P\,)V(\,I_m - Q\,), \tag{4.5}$$

$$X = PW + WQ - 2PWQ, \tag{4.6}$$

where $U,\ V,\ W \in \mathcal{C}^{m\times n}$ are arbitrary.

Proof. According to Corollary 4.2(c), the matrix equation $PX = XQ$ is equivalent to the pair of matrix equations

$$PX(\,I - Q\,) = 0 \quad \text{and} \quad (\,I - P\,)XQ = 0. \tag{4.7}$$

Solving the pair of equations, we can find that both (4.5) and (4.6) are the general solutions of $PX = XQ$. The process is somewhat tedious. Instead, we give here a direct verification. Putting (4.5) in PX and XQ, we get

$$PX = PUQ \quad \text{and} \quad XQ = PUQ.$$

Thus (4.5) is solution of $PX = XQ$. On the other hand, suppose that X_0 is a solution of $PX = XQ$ and let $U = V = X_0$ in (4.5). Then (4.5) becomes

$$X = PX_0Q + (\,I_m - P\,)X_0(\,I_m - Q\,) = PX_0Q + X_0 - PX_0 - X_0Q + PX_0Q = X_0,$$

which implies that any solution of $PX = XQ$ can be expressed by (4.5). Hence (4.5) is indeed the general solution of the equation $PX = XQ$. Similarly we can verify that (4.6) is also a general solution to $PX = XQ$. $\square$

As one of the basic linear matrix equation, $AX = XB$ was examined (see, e.g., Hartwig [56], Horn and Johnson [70], Parker [112], Slavova et al [125]). In general cases, solutions of $AX = XB$ can only be determined through canonical forms of A and B. The result in Corollary 4.4 manifests that for idempotent matrices A and B, the general solution of $AX = XB$ can directly be written in A and B. Obviously, the result in Corollary 4.4 is also valid for an operator equation of the form $AX = XB$ when both A and B are idempotent operators.

Theorem 4.5. *Let $A \in \mathcal{C}^{m\times n}$ be given, $P \in \mathcal{C}^{m\times m}$ and $Q \in \mathcal{C}^{n\times n}$ be two idempotent matrices. Then the sum $PA + AQ$ satisfies the rank equalities*

$$r(\,PA + AQ\,) = r\begin{bmatrix} PA & AQ \\ Q & 0 \end{bmatrix} - r(Q) = r\begin{bmatrix} AQ & P \\ PA & 0 \end{bmatrix} - r(P), \tag{4.8}$$

$$r(\,PA + AQ\,) = r\begin{bmatrix} AQ - PAQ \\ PA \end{bmatrix} = r[\,PA - PAQ,\ AQ\,]. \tag{4.9}$$

Proof. Let $M = \begin{bmatrix} P & 0 & PA \\ 0 & Q & Q \\ P & AQ & 0 \end{bmatrix}$. Then it is easy to see by block elementary operations of matrices that

$$r(M) = r\begin{bmatrix} P & 0 & 0 \\ 0 & Q & 0 \\ 0 & 0 & PA + AQ \end{bmatrix} = r(P) + r(Q) + r(\,PA + AQ\,).$$

On the other hand, note that $P^2 = P$ and $Q^2 = Q$. We also obtain by block elementary operations of matrices that

$$\begin{aligned} r(M) &= r\begin{bmatrix} P & -PAQ & PA \\ 0 & 0 & Q \\ P & AQ & 0 \end{bmatrix} \\ &= r\begin{bmatrix} 2P & 0 & PA \\ 0 & 0 & Q \\ P & AQ & 0 \end{bmatrix} = r\begin{bmatrix} 2P & 0 & 0 \\ 0 & 0 & Q \\ 0 & AQ & -\frac{1}{2}PA \end{bmatrix} = r(P) + r\begin{bmatrix} PA & AQ \\ Q & 0 \end{bmatrix}, \end{aligned}$$

and

$$\begin{aligned} r(M) &= r\begin{bmatrix} 0 & -PAQ & PA \\ 0 & Q & Q \\ P & AQ & 0 \end{bmatrix} \\ &= r\begin{bmatrix} 0 & 0 & PA \\ 0 & 2Q & Q \\ P & AQ & 0 \end{bmatrix} = r\begin{bmatrix} 0 & 0 & PA \\ 0 & 2Q & 0 \\ P & 0 & -\frac{1}{2}AQ \end{bmatrix} = r(Q) + r\begin{bmatrix} AQ & P \\ PA & 0 \end{bmatrix}. \end{aligned}$$

Combining the above three rank equalities for M yields (4.8). Consequently applying (1.2) and (1.3) to the two block matrices in (4.8) yields (4.9). □

Corollary 4.6. *Let $A \in \mathcal{C}^{m\times n}$ be given, $P \in \mathcal{C}^{m\times m}$ and $Q \in \mathcal{C}^{n\times n}$ be two idempotent matrices.*

(a) *If $PAQ = 0$, then $r(\,PA + AQ\,) = r(PA) + r(AQ)$, or equivalently $R(PA) \cap R(AQ) = \{0\}$ and $R[(PA)^*] \cap R[(AQ)^*] = \{0\}$.*

(b) *$PA + AQ = 0 \Leftrightarrow PA = 0$ and $AQ = 0$.*

(c) *The general solution of the matrix equation $PX + XQ = 0$ is $X = (\,I - P\,)U(\,I - Q\,)$, where $U \in \mathcal{C}^{m\times n}$ is arbitrary.*

Proof. If $PAQ = 0$, then $r(\,PA-AQ\,) = r(PA)+r(AQ)$ by Theorem 4.1(b). Consequently $r(\,PA+AQ\,) = r(PA) + r(AQ)$ by Lemma 1.4(d). The result in Part (b) follows from (4.9). According to (b), the equation $PX + XQ = 0$ is equivalent to the pair of matrix equations $PX = 0$ and $XQ = 0$. According to Rao and Mitra [118], and Mitra [101], the common general solution of the pair of equation is exactly $X = (\,I - P\,)U(\,I - Q\,)$, where $U \in \mathcal{C}^{m\times n}$ is arbitrary. □

Corollary 4.7. *Let $A,\ P,\ Q \in \mathcal{C}^{m\times m}$ be given with $P,\ Q$ being two idempotent matrices. Then the following five statements are equivalent:*

(a) *The sum $PA + AQ$ is nonsingular.*

(b) $r\begin{bmatrix} PA & AQ \\ Q & 0 \end{bmatrix} = m + r(Q).$

(c) $r\begin{bmatrix} AQ & P \\ PA & 0 \end{bmatrix} = m + r(P).$

(d) $r[\,PA,\ AQ\,] = m$ *and* $R\begin{bmatrix} (PA)^* \\ (AQ)^* \end{bmatrix} \cap R\begin{bmatrix} Q^* \\ 0 \end{bmatrix} = \{0\}.$

(e) $r\begin{bmatrix} AQ \\ PA \end{bmatrix} = m$ *and* $R\begin{bmatrix} AQ \\ PA \end{bmatrix} \cap R\begin{bmatrix} P \\ 0 \end{bmatrix} = \{0\}.$

Proof. Follows from (4.8). □

Theorem 4.8. *Let $A \in \mathcal{C}^{m\times n}$ be given, $P \in \mathcal{C}^{m\times m}$ and $Q \in \mathcal{C}^{n\times n}$ be two idempotent matrices. Then the rank of $A - PA - AQ$ satisfies the equalities*

$$r(\,A - PA - AQ\,) = r\begin{bmatrix} A & P \\ Q & 0 \end{bmatrix} + r(PAQ) - r(P) - r(Q) = r(\,A - PA - AQ + PAQ\,) + r(PAQ). \quad (4.10)$$

In particular,

(a) *$PA + AQ = A \Leftrightarrow (\,I - P\,)A(\,I - Q\,) = 0$ and $PAQ = 0$.*

(b) *The general solution of the matrix equation* $PX + XQ = X$ *is* $X = (I - P)UQ + V(I - Q)$, *where* $U, \ V \in \mathcal{C}^{m\times n}$ *are arbitrary.*

Proof. According to (4.1), we first find that

$$r(A - PA - AQ) = r[(I - P)A - AQ] = r\begin{bmatrix} (I - P)A \\ Q \end{bmatrix} + r[AQ, \ I - P] - r(I - P) - r(Q).$$

According to (1.2) and (1.3), we also get

$$r\begin{bmatrix} (I - P)A \\ Q \end{bmatrix} = r\begin{bmatrix} A & P \\ Q & 0 \end{bmatrix} - r(P), \quad \text{and} \quad r[AQ, \ I - P] = r(PAQ) + r(I - P).$$

Combining the above three yields the first equality in (4.10). Consequently applying (1.4) to the block matrix in it yields the second equality in (4.10). Part (a) is a direct consequence of (4.10), Part (a) follows from Corollary 4.4. □

If replacing P and Q in Theorem 4.5 by $I_m - P$ and $I_m - Q$, we can also obtain two rank equalities for $2A - PA - AQ$. For simplicity we omit them here.

Theorem 4.9. *Let* $A \in \mathcal{C}^{m\times n}$ *be given,* $P \in \mathcal{C}^{m\times m}$ *and* $Q \in \mathcal{C}^{n\times n}$ *be two idempotent matrices. Then the rank of* $A - PAQ$ *satisfies the equality*

$$r(A - PAQ) = r\begin{bmatrix} A & AQ & P \\ PA & 0 & 0 \\ Q & 0 & 0 \end{bmatrix} - r(P) - r(Q) = r\begin{bmatrix} (I - P)A(I - Q) & (I - P)AQ \\ PA(I - Q) & 0 \end{bmatrix}. \tag{4.11}$$

In particular,

(a) $PAQ = A \Leftrightarrow (I - P)A(I - Q) = 0, \ \ (I - P)AQ = 0$ *and* $PA(I - Q) = 0 \Leftrightarrow PA = A$ *and* $AQ = A$.

(b) *The general solution of the matrix equation* $PXQ = X$ *is* $X = PUQ$, *where* $U \in \mathcal{C}^{m\times n}$ *is arbitrary.*

Proof. Note that $P^2 = P$ and $Q^2 = Q$. It is easy to find that

$$\begin{aligned} r\begin{bmatrix} A & AQ & P \\ PA & 0 & 0 \\ Q & 0 & 0 \end{bmatrix} &= r\begin{bmatrix} A & 0 & P \\ 0 & -PAQ & -P \\ Q & -Q & 0 \end{bmatrix} \\ &= r\begin{bmatrix} A & 0 & P \\ -PAQ & 0 & -P \\ 0 & -Q & 0 \end{bmatrix} \\ &= r\begin{bmatrix} A - PAQ & 0 & 0 \\ 0 & 0 & -P \\ 0 & -Q & 0 \end{bmatrix} = r(A - PAQ) + r(P) + r(Q), \end{aligned}$$

as required for the first equality in (4.11). Consequently applying (1.4) to its left side yields the second one in (4.11). Part (a) is a direct consequence of (4.11), Part (b) can trivially be verified. □

Applying (4.1) to powers of difference of two idempotent matrices, we also find following several results.

Theorem 4.10. *Let* $P, \, Q \in \mathcal{C}^{m\times m}$ *be two idempotent matrices. Then*

(a) $(P - Q)^3$ *satisfies the two rank equalities*

$$r[(P - Q)^3] = r\begin{bmatrix} P - PQP \\ Q \end{bmatrix} + r[Q - QPQ, \ P] - r(P) - r(Q), \tag{4.12}$$

$$r[(P - Q)^3] = r[P - PQP - PQ + (PQ)^2] + r[Q - QPQ - PQ + (PQ)^2]. \tag{4.13}$$

In particular,

(b) *If* $(PQ)^2 = PQ$, *then*

$$r[(P - Q)^3] = r(P - PQP) + r(QPQ - Q). \tag{4.14}$$

(c) $r[\,(P-Q)^3\,] = r(\,P-Q\,)$, *i.e.*, $\mathrm{Ind}(\,P-Q\,) \leq 1$, *if and only if*

$$r\begin{bmatrix} P-PQP \\ Q \end{bmatrix} = r\begin{bmatrix} P \\ Q \end{bmatrix}, \quad \textit{and} \quad r[\,Q-QPQ,\ P\,] = r[\,Q,\ P\,], \tag{4.15}$$

or, equivalently,

$$R\left(\begin{bmatrix} P-PQP \\ Q \end{bmatrix}^*\right) = R\left(\begin{bmatrix} P \\ Q \end{bmatrix}^*\right), \quad \textit{and} \quad R[\,Q-QPQ,\ P\,] = R[\,Q,\ P\,]. \tag{4.16}$$

(d) $(\,P-Q\,)^3 = 0 \Leftrightarrow r\begin{bmatrix} P-PQP \\ Q \end{bmatrix} = r(Q)$ *and* $r[\,Q-QPQ,\ P\,] = r(P) \Leftrightarrow R(\,Q-QPQ\,) \subseteq R(P)$ *and* $R[(\,P-PQP\,)^*] \subseteq R(Q^*)$.

Proof. Since $P^2 = P$ and $Q^2 = Q$, it is easy to verify that

$$(\,P-Q\,)^3 = P(\,I_m - QP\,) - (\,I_m - QP\,)Q. \tag{4.17}$$

Letting $A = I_m - QP$ and applying (4.1) and (4.2) to (4.17) immediately yields (4.12) and (4.13). The results in Parts (b)—(d) are natural consequences of (4.13). □

Corollary 4.11. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then*

$$r[\,(\,P-Q\,)^3 - (\,P-Q\,)\,] = r\begin{bmatrix} PQP \\ Q \end{bmatrix} + r[\,QPQ,\ P\,] - r(P) - r(Q). \tag{4.18}$$

In particular,

(a) $P-Q$ *is tripotent* $\Leftrightarrow R(QPQ) \subseteq R(P)$ *and* $R[(PQP)^*] \subseteq R(Q^*)$.

(b) *If $PQ = QP$, then $P-Q$ is tripotent.*

Proof. Observe from (4.17) that

$$(\,P-Q\,)^3 - (\,P-Q\,) = -PQP + QPQ.$$

Applying (4.1) to it immediately yields (4.18). The results in Parts (b) and (c) are natural consequences of (4.18). □

Corollary 4.12. *A matrix $A \in \mathcal{C}^{m\times m}$ is tripotent if and only if it can factor as $A = P-Q$, where P and Q are two idempotent matrices with $PQ = QP$.*

Proof. The "if" part comes from Corollary 4.11(b). The " only if" part follows from a decomposition of A

$$A = \frac{1}{2}(\,A^2 + A\,) - \frac{1}{2}(\,A^2 - A\,),$$

where $P = \frac{1}{2}(\,A^2 + A\,)$ and $Q = \frac{1}{2}(\,A^2 - A\,)$ are two idempotent matrices with $PQ = QP$. □

The rank equality (4.12) can be extended to the matrix $(\,P-Q\,)^5$, where both P and Q are idempotent. In fact, it is easy to verify

$$(\,P-Q\,)^5 = P(\,I_m - QP\,)^2 - (\,I_m - QP\,)^2 Q.$$

Hence by (4.1) it follows that

$$r[\,(\,P-Q\,)^5\,] = r\begin{bmatrix} P(\,I_m - QP\,)^2 \\ Q \end{bmatrix} + r[\,(\,I_m - QP\,)^2 Q,\ P\,] - r(P) - r(Q).$$

Moreover, the above work can also be extended to $(\,P-Q\,)^{2k+1}(\,k = 3,\ 4,\ \cdots)$, where both P and Q are idempotent.

Applying (4.1) to $PQ - QP$, where both P and Q are idempotent, we also obtain the following.

Corollary 4.13. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then*

$$r(\,PQ - QP\,) = r\begin{bmatrix} PQ \\ P \end{bmatrix} + r[\,QP,\ P\,] - 2r(P), \tag{4.19}$$

$$r(\,PQ - QP\,) = r\begin{bmatrix} QP \\ Q \end{bmatrix} + r[\,PQ,\ Q\,] - 2r(Q), \tag{4.20}$$

$$r(\,PQ - QP\,) = r(\,PQ - PQP\,) + r(\,PQP - QP\,), \tag{4.21}$$

$$r(\,PQ - QP\,) = r(\,PQ - QPQ\,) + r(\,QPQ - QP\,). \tag{4.22}$$

In particular, if both P and Q are Hermitian idempotent, then

$$r(\,PQ - QP\,) = 2r(\,PQ - PQP\,) = 2r(\,PQ - QPQ\,). \tag{4.23}$$

The rank equality (4.23) was proved by Bérubé, Hartwig and Styan [17].

Corollary 4.14. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then*

$$r[\,(\,P - PQ\,) + \lambda(\,PQ - Q\,)] = r(\,P - Q\,) \tag{4.24}$$

holds for all $\lambda \in \mathcal{C}$ with $\lambda \neq 0$. In particular,

$$r(\,P + Q - 2PQ\,) = r(\,P + Q - 2QP\,) = r(\,P - Q\,). \tag{4.25}$$

Proof. Observe that

$$(\,P - PQ\,) + \lambda(\,PQ - Q\,) = P(\,P + \lambda Q\,) - (\,P + \lambda Q\,)Q.$$

Thus it follows by (4.1) that

$$\begin{aligned} r[\,(\,P - PQ\,) + \lambda(\,PQ - Q\,)] &= r\begin{bmatrix} P(\,P + \lambda Q\,) \\ Q \end{bmatrix} + r[\,(\,P + \lambda Q\,)Q,\ P\,] - r(P) - r(Q) \\ &= r\begin{bmatrix} P \\ Q \end{bmatrix} + r[\,\lambda Q,\ P\,] - r(P) - r(Q) \\ &= r\begin{bmatrix} P \\ Q \end{bmatrix} + r[\,P,\ Q\,] - r(P) - r(Q). \end{aligned}$$

Contrasting it with (3.1) yields (4.24). Setting $\lambda = -1$ we have (4.25). □

Replacing P by $I_m - P$ in (4.24), we also obtain the following.

Corollary 4.15. *Let $P,\ Q \in \mathcal{C}^{m\times m}$ be two idempotent matrices. Then*

$$r(\,I_m - P - Q + \lambda PQ\,) = r(\,I_m - P - Q\,)$$

holds for all $\lambda \in \mathcal{C}$ with $\lambda \neq 1$.

In the remainder of this chapter, we apply the results in Chapter 3 to establish various rank equalities related to involutory matrices. A matrix A is said to be involutory if its square is identity, i.e., $A^2 = I$. As two special types of matrices, involutory matrices and idempotent matrices are closely linked. As a matter of fact, for any involutory matrix A, the two corresponding matrices $(\,I + A\,)/2$ and $(\,I - A\,)/2$ are idempotent. Conversely, for any idempotent matrix A, the two corresponding matrices $\pm(\,I - 2A\,)$ are involutory. Based on the basic fact, all the results in Chapter 3 and this chapter on idempotent matrices can dually be extended to involutory matrices. We next list some of them.

Theorem 4.16. *Let $A,\ B \in \mathcal{C}^{m\times m}$ be two involutory matrices. Then the ranks of $A + B$ and $A - B$ satisfy the equalities*

$$r(\,A + B\,) = r\begin{bmatrix} I + A \\ I - B \end{bmatrix} + r[\,I + A,\ I - B\,] - r(\,I + A\,) - r(\,I - B\,), \tag{4.26}$$

$$r(\,A + B\,) = r[\,(\,I + A\,)(\,I + B\,)\,] + r[\,(\,I - A\,)(\,I - B\,)\,], \tag{4.27}$$

$$r(\,A - B\,) = r\begin{bmatrix} I + A \\ I + B \end{bmatrix} + r[\,I + A,\ I + B\,] - r(\,I + A\,) - r(\,I + B\,), \tag{4.28}$$

$$r(\,A - B\,) = r[\,(\,I + A\,)(\,I - B\,)\,] + r[\,(\,I - A\,)(\,I + B\,)\,]. \tag{4.29}$$

Proof. Notice that both $P = (I+A)/2$ and $Q = (I-B)/2$ are idempotent when A and B are involutory. In that case,

$$r(P-Q) = r\left[\frac{1}{2}(I+A) - \frac{1}{2}(I-B)\right] = r(A+B),$$

and

$$r\begin{bmatrix} P \\ Q \end{bmatrix} + r[P,\ Q] - r(P) - r(Q) = r\begin{bmatrix} I+A \\ I-B \end{bmatrix} + r[I+A,\ I-B] - r(I+A) - r(I-B).$$

Putting them in (3.1) produces (4.26). Furthermore we have

$$r(P-PQ) = r\left[(I+A)\left(I - \frac{1}{2}(I-B)\right)\right] = r[(I+A)(I+B)],$$
$$r(PQ-Q) = r\left[\left(\frac{1}{2}(I+A) - I\right)(I-B)\right] = r[(I-A)(I-B)].$$

Putting them in (3.2) yields (4.27). moreover, if B is involutory, then $-B$ is also involutory. Thus replacing B by $-B$ in (4.26) and (4.27) yields (4.28) and (4.29). □

Corollary 4.17. *Let $A,\ B \in \mathcal{C}^{m\times m}$ be two involutory matrices.*

(a) *If $(I+A)(I-B) = 0$ or $(I-B)(I+A) = 0$, then*

$$r(A+B) = r(I+A) + r(I-B). \tag{4.30}$$

(b) *If $(I+A)(I+B) = 0$ or $(I+B)(I+A) = 0$, then*

$$r(A-B) = r(I+A) + r(I+B). \tag{4.31}$$

Proof. The condition $(I+A)(I-B) = 0$ is equivalent to $I+A = B+BA$ and $I-B = AB-A$. In that case, $(I+A)(I+B) = I+A+B+AB = 2(I+A)$. and $(I-A)(I-B) = I-B-A+AB = 2(I-B)$. Thus (4.27) reduces to (4.30). Similarly we show that under $(I-B)(I+A) = 0$, the rank equality (4.30) also holds. The result in Part (b) is obtained by replacing B in Part (a) by $-B$. □

Corollary 4.18. *Let $A,\ B \in \mathcal{C}^{m\times m}$ be two involutory matrices. Then*

(a) *The sum $A+B$ is nonsingular if and only if*

$$R(I+A)\cap R(I-B) = \{0\}, \quad R(I+A^*)\cap R(I-B^*) = \{0\}, \quad \text{and} \quad r(I+A) + r(I-B) = m.$$

(b) *The difference $A-B$ is nonsingular if and only if*

$$R(I+A)\cap R(I+B) = \{0\}, \quad R(I+A^*)\cap R(I+B^*) = \{0\}, \quad \text{and} \quad r(I+A) + r(I+B) = m.$$

Proof. Follows immediately from (4.26) and (4.27). □

Theorem 4.19. *Let $A,\ B \in \mathcal{C}^{m\times m}$ be two involutory matrices. Then $A+B$ and $A-B$ satisfy the rank equalities*

$$r(A+B) = r[(I+A)(I+B)] + r[(I+B)(I+A)] - r(I+A) - r(I+B) + m, \tag{4.32}$$
$$r(A-B) = r[(I+A)(I-B)] + r[(I-B)(I+A)] - r(I+A) - r(I-B) + m. \tag{4.33}$$

Proof. Putting $P = (I+A)/2$ and $Q = (I+B)/2$ in (3.8) and simplifying yields (4.32). Replacing B by $-B$ in (4.32) yields (4.33). □

The combination of (4.27) with (4.32) produces the following rank equality

$$r[(I+B)(I+A)] = r(I+B) + r(I+A) - m + r[(I-A)(I-B)]. \tag{4.34}$$

Theorem 4.20. *Let $A,\ B \in \mathcal{C}^{m\times m}$ be two involutory matrices. Then*

$$r(AB-BA) = r(A+B) + r(A-B) - m. \tag{4.35}$$

Inparticular,

$$AB = BA \Leftrightarrow r(A+B) + r(A-B) = m. \tag{4.36}$$

Proof. Putting $P = (I + A)/2$ and $Q = (I - B)/2$ in (3.24) and simplifying yields (4.35). $\square$

Putting the formulas (4.26)—(4.29), (4.32) and (4.33) in (4.35) may yield some other rank equalities for $AB - BA$. We leave them to the reader.

Theorem 4.21. *Let $A, B \in \mathcal{C}^{m\times m}$ be two involutory matrices. Then*
(a) $r\left[\left(\frac{A+B}{2}\right)^2 - \frac{A+B}{2}\right] = r(I - A - B) + r(A + B) - m.$
(b) $r\left[\left(\frac{A-B}{2}\right)^2 - \frac{A-B}{2}\right] = r(I - A + B) + r(A - B) - m.$
In particular,
(c) $\frac{1}{2}(A + B)$ *is idempotent* $\Leftrightarrow r(I - A - B) + r(A + B) = m \Leftrightarrow r(A + B) = r(I + A) - r(I - B).$
(d) $\frac{1}{2}(A - B)$ *is idempotent* $\Leftrightarrow r(I - A + B) + r(A - B) = m \Leftrightarrow r(A - B) = r(I + A) - r(I + B).$

Proof. Putting $P = (I + A)/2$ and $Q = (I - B)/2$ in (3.32) and simplifying yields Part (a). Replacing B by $-B$ we get Part (b). Part (c) and (d) follow from Parts (a) and (b), and Corollary 3.21(c). $\square$

Theorem 4.22. *Let $A, B \in \mathcal{C}^{m\times m}$ be two involutory matrices. Then*

$$r(3I - A - B - AB) = r(2I - A - B). \tag{4.39}$$

Proof. Putting $P = (I + A)/2$ and $Q = (I + B)/2$ in (3.34) and simplifying yields (4.39). $\square$

Theorem 4.23. *Let $A \in \mathcal{C}^{m\times m}$ be an involutory matrix. Then*
(a) $r(A - A^*) = 2r[I + A,\ I + A^*] - 2r(I + A) = r[I - A,\ I - A^*] - 2r(I - A).$
(b) $r(A + A^*) = m.$
(c) $r(AA^* - A^*A) = r(A - A^*).$

Proof. Putting $P = (I \pm A)/2$ and $Q = (I \pm A^*)/2$ in Corollary 3.26 and simplifying yields the desired results. $\square$

Theorem 4.24. *Let $A \in \mathcal{C}^{m\times m}$ and $B \in \mathcal{C}^{n\times n}$ be two involutory matrices, and $X \in \mathcal{C}^{m\times m}$. Then $AX - XB$ satisfies the rank equalities*

$$r(AX - XB) = r\begin{bmatrix} (I_m + A)X \\ I_n + B \end{bmatrix} + r[X(I_n + B),\ I_m + A] - r(I_m + A) - r(I_n + B), \tag{4.40}$$

$$r(AX - XB) = r[(I_m + A)X(I_n - B)] + r[(I_m - A)X(I_n + B)]. \tag{4.41}$$

In particular,

$$AX = XB \Leftrightarrow (I_m + A)X(I_n - B) = 0 \ \text{ and } \ (I_m - A)X(I_n + B) = 0. \tag{4.42}$$

Proof. Putting $P = (I_m + A)/2$ and $Q = (I_n + B)/2$ in (4.1) and (4.2) yields (4.40) and (4.41). The equivalence in (4.42) follows from (4.41). $\square$

Theorem 4.25. *Let $A \in \mathcal{C}^{m\times m}$ and $B \in \mathcal{C}^{n\times n}$ be two involutory matrices. Then the general solution of the matrix equation $AX = XB$ is*

$$X = V + AVB, \tag{4.43}$$

where $V \in \mathcal{C}^{m\times n}$ is arbitrary.

Proof. We only give the verification. Obviously the matrix X in (4.43) satisfies $AX = AV + VB$ and $XB = VB + AV$. Thus X is a solution of $AX = XB$. On the other hand, for any solution X_0 of $AX = XB$, let $V = X_0/2$ in (4.43). Then we get $V = AX_0B = X_0$, that is, X_0 can be represented by (4.43). Thus (4.43) is the general solution of the matrix equation $AX = XB$. $\square$

Theorem 4.26. *Let $A \in \mathcal{C}^{m\times m}$ be an involutory matrix, and $X \in \mathcal{C}^{m\times m}$. Then*
(a) *$AX - XA$ satisfies the rank equalities*

$$r(AX - XA) = r\begin{bmatrix} (I + A)X \\ I + A \end{bmatrix} + r[X(I + A),\ I + A] - 2r(I + A),$$

$$r(AX - XA) = r[(I + A)X(I - A)] + r[(I - A)X(I + A)],$$

In particular,

$$AX = XA \Leftrightarrow (I + A)X(I - A) = 0 \ \text{ and } \ (I - A)X(I + A) = 0.$$

(b) *The general solution of the matrix equation* $AX = XA$ *is*

$$X = V + AVA,$$

where $V \in \mathcal{C}^{m \times m}$ *is arbitrary.*

Chapter 5

Rank equalities for outer inverses of matrices

An outer inverse of a matrix A is the solution to the matrix equation $XAX = X$, and is often denoted by $X = A^{(2)}$. The collection of all outer inverses of A is often denoted by $A\{2\}$. Obviously, the Moore-Penrose inverse, the Drazin inverse, the group inverse, and the weighted Moore-Penrose inverse of a matrix are naturally outer inverses of the matrix. If outer inverse of a matrix is also an inner inverse the matrix, it is called a reflexive inner inverse of the matrix, and is often denoted by A_r^-. The collection of all reflexive inner inverses of a matrix A is denoted by $A\{1, 2\}$. As one of important kinds of generalized inverses of matrices, outer inverses of matrices and their applications have well been examined in the literature (see, e.g., [16, 21, 46, 71, 108, 147, 148]). In this chapter, we shall establish several basic rank equalities related to differences and sums of outer inverses of a matrix, and then consider their various consequences. The results obtained in this chapter will also be applied in the subsequent chapters.

Theorem 5.1. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1, X_2 \in A\{2\}$. Then the difference $X_1 - X_2$ satisfies the following three rank equalities*

$$r(X_1 - X_2) = r\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + r[\, X_1, \ X_2\,] - r(X_1) - r(X_2), \tag{5.1}$$

$$r(X_1 - X_2) = r(X_1 - X_1AX_2) + r(X_1AX_2 - X_2), \tag{5.2}$$

$$r(X_1 - X_2) = r(X_1 - X_2AX_1) + r(X_2AX_1 - X_2). \tag{5.3}$$

Proof. Let $M = \begin{bmatrix} -X_1 & 0 & X_1 \\ 0 & X_2 & X_2 \\ X_1 & X_2 & 0 \end{bmatrix}$. Then it is easy to see by block elementary operations of matrices that

$$r(M) = r\begin{bmatrix} -X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & X_1 - X_2 \end{bmatrix} = r(X_1) + r(X_2) + r(X_1 - X_2). \tag{5.4}$$

On the other hand, note that $X_1AX_1 = X_1$ and $X_2AX_2 = X_2$. Thus

$$\begin{bmatrix} I_n & 0 & X_1A \\ 0 & I_n & 0 \\ 0 & 0 & I_n \end{bmatrix}\begin{bmatrix} -X_1 & 0 & X_1 \\ 0 & X_2 & X_2 \\ X_1 & X_2 & 0 \end{bmatrix}\begin{bmatrix} I_m & 0 & 0 \\ 0 & I_m & 0 \\ 0 & -AX_2 & I_m \end{bmatrix} = \begin{bmatrix} 0 & 0 & X_1 \\ 0 & 0 & X_2 \\ X_1 & X_2 & 0 \end{bmatrix},$$

which implies that

$$r(M) = r\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + r[\, X_1, \ X_2\,]. \tag{5.5}$$

Combining (5.4) and (5.5) yields (5.1). Consequently applying (1.2) and (1.3) to the two block matrices in (5.1) respectively and noticing that $A \in X_1\{2\}$ and $A \in X_2\{2\}$, we can write (5.1) as (5.2) and (5.3). $\square$

It is obvious that if $A = I_m$ in Theorem 5.1, then $X_1,\ X_2 \in I_m\{2\}$ are actually two idempotent matrices. In that case, (5.1)—(5.3) reduce to the results in Theorem 3.1.

Corollary 5.2. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1,\ X_2 \in A\{2\}$. Then*

(a) $R(X_1 - X_1AX_2) \cap R(X_1AX_2 - X_2) = \{0\}$ *and* $R[(X_1 - X_1AX_2)^*] \cap R[(X_1AX_2 - X_2)^*] = \{0\}$.

(b) $R(X_1 - X_2AX_1) \cap R(X_2AX_1 - X_2) = \{0\}$ *and* $R[(X_1 - X_2AX_1)^*] \cap R[(X_2AX_1 - X_2)^*] = \{0\}$.

(c) *If* $X_1AX_2 = 0$ *or* $X_2AX_1 = 0$, *then* $r(X_1 - X_2) = r(X_1) + r(X_2)$.

Proof. The results in Parts (a) and (b) follow immediately from applying Lemma 1.4(d) to (5.2) and (5.3). Parts (c) is a direct consequence of (5.2) and (5.3). □

Corollary 5.3. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1,\ X_2 \in A\{2\}$. Then the following five statements are equivalent*:

(a) $r(X_1 - X_2) = r(X_1) - r(X_2)$, *i.e.*, $X_2 \leq_{rs} X_1$.

(b) $r\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = r[\, X_1,\ X_2\,] = r(X_1)$.

(c) $R(X_2) \subseteq R(X_1)$ *and* $R(X_2^*) \subseteq R(X_1^*)$.

(d) $X_1AX_2 = X_2$ *and* $X_2AX_1 = X_2$.

(e) $X_1AX_2AX_1 = X_2$.

Proof. The equivalence of Parts (a) and (b) follows directly from (5.1). The equivalence of Parts (b), (c) and (d) follows directly from Lemma 1.2(c) and (d). Combining the two equalities in Part (d) yields the equality in Part (e). Conversely, suppose that $X_1AX_2AX_1 = X_2$ holds. Pre- and post-multiplying X_1A and AX_1 to it yields $X_1AX_2AX_1 = X_1AX_2 = X_2AX_1$. Combining it with $X_1AX_2AX_1 = X_2$ yields the two rank equalities in Part (d). □

Corollary 5.4. *Let $A \in \mathcal{C}^{m\times m}$ be given, and $X_1,\ X_2 \in A\{2\}$. Then the following three statements are equivalent*:

(a) *The difference $X_1 - X_2$ is nonsingular.*

(b) $r\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = r[\, X_1,\ X_2\,] = r(X_1) + r(X_2) = m$.

(c) $R(X_1) \oplus R(X_2) = R(X_1^*) \oplus R(X_2^*) = \mathcal{C}^m$.

Proof. A trivial consequence of (5.1). □

Corollary 5.5. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X \in A\{2\}$. Then*

$$r(A - AXA) = r(A) - r(AXA), \quad i.e., \quad AXA \leq_{rs} A. \tag{5.6}$$

In particular,

$$AXA = A, \ i.e., \ X \in A\{1,\, 2\} \Leftrightarrow r(A) = r(X). \tag{5.7}$$

Proof. It is easy to verify that both A and AXA are outer inverses of $A^\dagger$. Thus by (5.1) we obtain

$$r(A - AXA) = r\begin{bmatrix} A \\ AXA \end{bmatrix} + r[\, A,\ AXA\,] - r(A) - r(AXA) = r(A) - r(AXA),$$

the desired in (5.6). □

Corollary 5.6. *Let $A \in \mathcal{C}^{m\times m}$ be given, and $X \in A\{2\}$. Then*

$$r(AX - XA) = r\begin{bmatrix} X \\ XA \end{bmatrix} + r[\, X,\ AX\,] - 2r(X) = r(XA - XA^2X) + r(XA^2X - AX). \tag{5.8}$$

Inparticular,

$$AX = XA \Longleftrightarrow R(AX) = R(X) \ \ and \ \ R[(AX)^*] = R(X^*). \tag{5.9}$$

Proof. It is easy to verify that both AX and XA are idempotent when $X \in A\{2\}$. Thus we find by (3.1), (1.2) and (1.3) that

$$r(AX - XA) \ \ = \ \ r\begin{bmatrix} AX \\ XA \end{bmatrix} + r[\, AX,\ XA\,] - r(AX) - r(XA)$$

$$= r\begin{bmatrix} X \\ XA \end{bmatrix} + r[\,X,\ AX\,] - r(AX) - r(XA)$$
$$= r(\,XA - XA^2X\,) + r(\,XA^2X - AX\,),$$

as required for (5.8). Eq. (5.9) is a direct consequence of (5.8). □

Corollary 5.7. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1,\, X_2 \in A\{2\}$. Then*

$$r(\,AX_1A - AX_2A\,) = r\begin{bmatrix} X_1A \\ X_2A \end{bmatrix} + r[\,AX_1,\ AX_2\,] - r(X_1) - r(X_2). \tag{5.10}$$

In particular,

$$AX_1A = AX_2A \Leftrightarrow X_1AX_2AX_1 = X_1 \text{ and } X_2AX_1AX_2 = X_2. \tag{5.11}$$

Proof. Notice that Both AX_1A and AX_2A are outer inverses of $A^\dagger$ when $X_1,\ X_2 \in A\{2\}$. Moreover, observe that $r(AX_1A) = r(AX_1) = r(X_1A) = r(X_1)$, and $r(AX_2A) = r(AX_2) = r(X_2A) = r(X_2)$. Thus it follows from (5.1) that

$$r(\,AX_1A - AX_2A\,) = r\begin{bmatrix} AX_1A \\ AX_2A \end{bmatrix} + r[\,AX_1A,\ AX_2A\,] - r(AX_1A) - r(AX_2A)$$
$$= r\begin{bmatrix} X_1A \\ X_2A \end{bmatrix} + r[\,AX_1,\ AX_2\,] - r(X_1) - r(X_2),$$

as required for (5.10). The verification of (5.11) is trivial, hence is omitted. □

Corollary 5.8. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1,\, X_2 \in A\{2\}$. Then the following five statements are equivalent*:

(a) $r(\,AX_1A - AX_2A\,) = r(AX_1A) - r(AX_2A)$, *i.e.*, $AX_2A \leq_{rs} AX_1A$.

(b) $\begin{bmatrix} X_1A \\ X_2A \end{bmatrix} = r[\,AX_1,\ AX_2\,] = r(X_1)$.

(c) $R(AX_2) \subseteq R(AX_1)$ *and* $R[(X_2A)^*] \subseteq R[(X_1A)^*]$.

(d) $AX_1AX_2A = AX_2A$ *and* $AX_2AX_1A = AX_2A$.

(e) $AX_1AX_2AX_1A = AX_2A$.

Proof. Follows form Corollary 5.3 by noticing that Both AX_1A and AX_2A are outer inverses of $A^\dagger$ when $X_1,\ X_2 \in A\{2\}$. □

Theorem 5.9. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1,\, X_2 \in A\{2\}$. Then the sum $X_1 + X_2$ satisfies the rank equalities*

$$r(\,X_1 + X_2\,) = r\begin{bmatrix} X_1 & X_2 \\ X_2 & 0 \end{bmatrix} - r(X_2) = r\begin{bmatrix} X_2 & X_1 \\ X_1 & 0 \end{bmatrix} - r(X_1), \tag{5.11}$$
$$r(\,X_1 + X_2\,) = r[\,(\,I_n - X_2A\,)X_1(\,I_m - AX_2\,)\,] + r(X_2), \tag{5.12}$$
$$r(\,X_1 + X_2\,) = r[\,(\,I_n - X_1A\,)X_2(\,I_m - AX_1\,)\,] + r(X_1). \tag{5.13}$$

Proof. Let $M = \begin{bmatrix} X_1 & 0 & X_1 \\ 0 & X_2 & X_2 \\ X_1 & X_2 & 0 \end{bmatrix}$. Then it is easy to see by block elementary operations that

$$r(M) = r\begin{bmatrix} X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & -(\,X_1 + X_2\,) \end{bmatrix} = r(X_1) + r(X_2) + r(\,X_1 + X_2\,). \tag{5.14}$$

On the other hand, note that $X_1AX_1 = X_1$ and $X_2AX_2 = X_2$. Thus

$$\begin{bmatrix} I_n & 0 & X_1A \\ 0 & I_n & 0 \\ 0 & 0 & I_n \end{bmatrix}\begin{bmatrix} X_1 & 0 & X_1 \\ 0 & X_2 & X_2 \\ X_1 & X_2 & 0 \end{bmatrix}\begin{bmatrix} I_m & 0 & 0 \\ 0 & I_m & 0 \\ 0 & -AX_2 & I_m \end{bmatrix} = \begin{bmatrix} 2X_1 & 0 & X_1 \\ 0 & 0 & X_2 \\ X_1 & X_2 & 0 \end{bmatrix},$$

which implies that

$$r(M) = r\begin{bmatrix} 2X_1 & 0 & X_1 \\ 0 & 0 & X_2 \\ X_1 & X_2 & 0 \end{bmatrix} = r\begin{bmatrix} 2X_1 & 0 & 0 \\ 0 & 0 & X_2 \\ 0 & X_2 & -\frac{1}{2}X_1 \end{bmatrix} = r\begin{bmatrix} X_1 & X_2 \\ X_2 & 0 \end{bmatrix} + r(X_1). \tag{5.15}$$

Combining (5.14) and (5.15) yields the first equality in (5.11). By symmetry, we have the second equality in (5.15). Applying (1.3) to the two block matrices in (5.11), respectively, and noticing that $A \in \{X_1^-\}$ and $A \in \{X_2^-\}$, we then can write (5.11) as (5.12) and (5.13). □

Corollary 5.10. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1,\ X_2 \in A\{2\}$.*

(a) *If $X_1AX_2 = X_2AX_1$, then $r(\, X_1 + X_2\,) = r\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = r[\, X_1,\ X_2\,]$.*

(b) *If $X_1AX_2 = X_2AX_1 = 0$, then $r(\, X_1 + X_2\,) = r(X_1) + r(X_2)$.*

Proof. Under $X_1AX_2 = X_2AX_1$, we find from (5.12) and (5.13) that

$$r(\, X_1 + X_2\,) = r(\, X_1 - X_1AX_2\,) + r(X_2) = r(X_1) + r(\, X_1AX_2 - X_2\,).$$

Note by (1.2) and (1.3) that

$$r\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = r(\, X_1 - X_1AX_2\,) + r(X_2), \quad \text{and} \quad r[\, X_1,\ X_2\,] = r(X_1) + r(\, X_1AX_2 - X_2\,).$$

Thus we have the results in Part (a). Part (b) follows immediately from (5.12). □

Corollary 5.11. *Let $A \in \mathcal{C}^{m\times m}$ be given, and $X_1,\ X_2 \in A\{2\}$. Then the following five statements are equivalent*:

(a) *The sum $X_1 + X_2$ is nonsingular.*

(b) $r\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = m$ *and* $R\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \cap R\begin{bmatrix} X_2 \\ 0 \end{bmatrix} = \{0\}$.

(c) $r[\, X_1,\ X_2\,] = m$ *and* $R\begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix} \cap R\begin{bmatrix} X_2^* \\ 0 \end{bmatrix} = \{0\}$.

(d) $r\begin{bmatrix} X_2 \\ X_1 \end{bmatrix} = m$ *and* $R\begin{bmatrix} X_2 \\ X_1 \end{bmatrix} \cap R\begin{bmatrix} X_1 \\ 0 \end{bmatrix} = \{0\}$.

(e) $r[\, X_2,\ X_1\,] = m$ *and* $R\begin{bmatrix} X_2^* \\ X_1^* \end{bmatrix} \cap R\begin{bmatrix} X_1^* \\ 0 \end{bmatrix} = \{0\}$.

Proof. Follows immediately from (5.11). □

Corollary 5.12. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X \in A\{2\}$. Then*

$$r(\, A + AXA\,) = r(A). \tag{5.16}$$

holds for all $X \in A\{2\}$.

Proof. Notice that Both A and AX_2A are outer inverses of $A^{\dagger}$ when $X \in A\{2\}$. Thus (5.16) follows from (5.11). □

Theorem 5.13. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1,\ X_2 \in A\{2\}$. Then the difference $X_1 - X_2$ satisfies the rank equalities*

$$r[\,(\, X_1 - X_2\,)A(\, X_1 - X_2\,) - (\, X_1 - X_2\,)\,] = r(\, I_m - AX_1 + AX_2\,) + r(\, X_1 - X_2\,) - m, \tag{5.17}$$

$$r[\,(\, X_1 - X_2\,)A(\, X_1 - X_2\,) - (\, X_1 - X_2\,)\,] = r(X_1AX_2AX_1\,) - r(X_1) + r(\, X_1 - X_2\,). \tag{5.18}$$

Proof. Letting $X = X_1 - X_2$ and applying (1.10) yields

$$r(\, XAX - X\,) = r(\, I_m - AX\,) + r(X) - m,$$

which is (5.17). Note that AX_1 and AX_2 are idempotent. It turns out by (3.19) that

$$r(\, I_m - AX_1 + AX_2\,) = r(AX_1AX_2AX_1) - r(AX_1) + m = r(X_1AX_2AX_1) - r(X_1) + m.$$

Putting it in (5.17) yields (5.18). □

Corollary 5.14. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1,\, X_2 \in A\{2\}$. Then the following five statements are equivalent*:

(a) $X_1 - X_2 \in A\{2\}$.

(b) $r(\, I_m - AX_1 + AX_2\,) = m - r(\, X_1 - X_2\,)$.

(c) $r(\, X_1 - X_2\,) = r(X_1) - r(X_2)$, *i.e.*, $X_2 \leq_{rs} X_1$.

(d) $R(X_2) \subseteq R(X_1)$ *and* $R(X_2^*) \subseteq R(X_1^*)$.

(e) $X_1AX_2AX_1 = X_2$.

Proof. The equivalence of Parts (a) and (b) follows immediately from (5.17). The equivalence of Parts (c), (d) and (e) is from Corollary 5.3. We next show the equivalence of Parts (a) and (e). It is easy to verify that

$$(\, X_1 - X_2\,)A(\, X_1 - X_2\,) - (\, X_1 - X_2\,) = -X_1AX_2 - X_2AX_1 + 2X_2,$$

Thus $X_1 - X_2 \in A\{2\}$ holds if and only if

$$X_1AX_2 + X_2AX_1 = 2X_2. \tag{5.19}$$

Pre- and post-multiplying X_1A and AX_1 to it, we get

$$X_1AX_2AX_1 = X_1AX_2 \quad \text{and} \quad X_1AX_2AX_1 = X_2AX_1. \tag{5.20}$$

Putting them in (5.19) yields Part (e). Conversely, if Part (e) holds, then (5.20) holds. Combining Part (e) with (5.20) leads to (5.19), which is equivalent to $X_1 - X_2 \in A\{2\}$. □

The problem considered in Corollary 5.14 could be regarded as an extension of the work in Corollary 3.21, which was examined by Getson and Hsuan citeGH. In that monograph, they only gave a sufficient condition for $X_1 - X_2 \in A\{2\}$ to hold when $X_1,\ X_2 \in A\{2\}$. Our result in Corollary 5.14 is a complete conclusion on this problem.

Theorem 5.15. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1,\, X_2 \in A\{2\}$. Then the sum $X_1 + X_2$ satisfies the two rank equalities*

$$r[\,(\, X_1 + X_2\,)A(\, X_1 + X_2\,) - (\, X_1 + X_2\,)\,] = r(\, I_m - AX_1 - AX_2\,) + r(\, X_1 + X_2\,) - m, \tag{5.21}$$

$$r[\,(\, X_1 + X_2\,)A(\, X_1 + X_2\,) - (\, X_1 + X_2\,)\,] = r(X_1AX_2\,) + r(X_2AX_1\,) + r(\, X_1 + X_2\,) - r(X_1) - r(X_2). \tag{5.22}$$

Proof. Letting $X = X_1 + X_2$ and applying (1.10) to $XAX - X$ yields (5.21). Note that AX_1 and AX_2 are idempotent. It turns out by (3.8) that

$$r(\, I_m - AX_1 - AX_2\,) = r(X_1AX_2\,) + r(X_2AX_1\,) - r(X_1) - r(X_2) + m.$$

Putting it in (5.21) yields (5.22). □

Corollary 5.16. *Let $A \in \mathcal{C}^{m\times n}$ be given, and $X_1,\, X_2 \in A\{2\}$. Then the following four statements are equivalent*:

(a) $X_1 + X_2 \in A\{2\}$.

(b) $X_1AX_2 + X_2AX_1 = 0$.

(c) $r(\, I_m - AX_1 - AX_2\,) = m - r(\, X_1 + X_2\,)$.

(d) $X_1AX_2 = 0$, *and* $X_2AX_1 = 0$.

Proof. The equivalence of Parts (a) and (b) follows immediately from expanding $(\, X_1 + X_2\,)A(\, X_1 + X_2\,) - (\, X_1 + X_2\,)$. The equivalence of (a) and (c) is from (5.21). We next show the equivalence of (b) and (d). Pre- and post-multiplying X_1A and AX_1 to $X_1AX_2 + X_2AX_1 = 0$, we get

$$X_1AX_2 + X_1AX_2AX_1 = 0, \quad \text{and} \quad X_1AX_2AX_1 + X_2AX_1 = 0,$$

which implies that $X_1AX_2 = X_2AX_1$. Putting them in Part (b) yields (d). Conversely, if Part (d) holds, then Part (b) naturally holds. □

Chapter 6

Rank equalities for a matrix and its Moore-Penrose inverse

In this chapter, we shall establish a variety of rank equalities related to a matrix and its Moore-Penrose inverse, and then use them to characterize various specified matrices, such as, EP matrices, conjugate EP matrices, bi-EP matrices, star-dagger matrices, and so on.

A matrix A is said to be EP (or Range-Hermitian) if $R(A) = R(A^*)$. EP matrices have some nice properties, meanwhile they are quite inclusive. Hermitian matrices, normal matrices, as well as nonsingular matrices are special cases of EP matrices. As a class of important matrices, EP matrices and their applications have well be examined in the literature. One of the basic and nice properties related to an EP matrix A is $AA^{\dagger} = A^{\dagger}A$, see, e.g., Ben-Israel and Greville [16], Campbell and Meyer [21]. This equality motivates us to consider the rank of $AA^{\dagger} - A^{\dagger}A$, as well as its various extensions.

Theorem 6.1. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then the rank of $AA^{\dagger} - A^{\dagger}A$ satisfies the following rank equalities*

$$r(\, AA^{\dagger} - A^{\dagger}A\,) = 2r[\, A, \ A^*\,] - 2r(A) = 2r(\, A - A^2A^{\dagger}\,) = 2r(\, A - A^{\dagger}A^2\,). \tag{6.1}$$

In particular,

(a) $AA^{\dagger} = A^{\dagger}A \Leftrightarrow r[\, A, \ A^*\,] = r(A) \Leftrightarrow A = A^2A^{\dagger} \Leftrightarrow A = A^{\dagger}A^2 \Leftrightarrow R(A) = R(A^*), \ i.e., \ A \ is \ EP.$

(b) $AA^{\dagger} - A^{\dagger}A$ *is nonsingular* $\Leftrightarrow r[\, A, \ A^*\,] = 2r(A) = m \Leftrightarrow R(A) \oplus R(A^*) = \mathcal{C}^m.$

Proof. Note that $AA^{\dagger}$ and $A^{\dagger}A$ are idempotent matrices. Then applying (3.1), we first obtain

$$r(\, AA^{\dagger} - A^{\dagger}A\,) = r\begin{bmatrix} AA^{\dagger} \\ A^{\dagger}A \end{bmatrix} + r[\, AA^{\dagger}, \ A^{\dagger}A\,] - r(AA^{\dagger}) - r(A^{\dagger}A). \tag{6.2}$$

Observe that $r(AA^{\dagger}) = r(A^{\dagger}A) = r(A)$, and

$$r\begin{bmatrix} AA^{\dagger} \\ A^{\dagger}A \end{bmatrix} = r\begin{bmatrix} A^{\dagger} \\ A \end{bmatrix} = r\begin{bmatrix} A^* \\ A \end{bmatrix}, \qquad r[\, AA^{\dagger}, \ A^{\dagger}A\,] = r[\, A, \ A^{\dagger}\,] = r[\, A, \ A^*\,].$$

Thus (6.2) reduces to the first rank equality in (6.1). Consequently applying (1.2) to $[\, A, \ A^*\,]$ in (6.1) yields the other two rank equalities in (6.1). The equivalence in Part (a) are well-known results on a EP matrix, which now is a direct consequence of (6.1). It remains to show Part (b). If $r[\, AA^{\dagger} - A^{\dagger}A\,] = m$, then $r[\, A, \ A^*\,] = r[\, AA^{\dagger}, \ A^{\dagger}A\,] = r[\, AA^{\dagger} - A^{\dagger}A, \ A^{\dagger}A\,] = m$. Putting it in (6.1), we obtain $2r(A) = m$. Conversely, if $r[\, A, \ A^*\,] = 2r(A) = m$, then we immediately have $r(\, AA^{\dagger} - A^{\dagger}A\,) = m$ by (6.1). Hence the first equivalence in Part (b) is true. The second equivalence is obvious. $\square$

Another group of rank equalities related to EP matrix is given below, which is motivated by a work of Campbell and Meyer [20].

Theorem 6.2. *Let $A \in \mathcal{C}^{m\times m}$ and $0 \neq k \in \mathcal{C}$ be given. Then*

(a) $r[\, AA^{\dagger}(\, A + kA^{\dagger}\,) - (\, A + kA^{\dagger}\,)AA^{\dagger}\,] = 2r[\, A, \ A^*\,] - 2r(A).$

(b) $r[\, A^{\dagger}A(\, A + kA^{\dagger}\,) - (\, A + kA^{\dagger}\,)A^{\dagger}A\,] = 2r[\, A, \ A^*\,] - 2r(A).$

(c) $r[\, AA^{\dagger}(\, A + kA^* \,) - (\, A + kA^* \,)AA^{\dagger} \,] = 2r[\, A, \;\; A^* \,] - 2r(A)$.
(d) $r[\, A^{\dagger}A(\, A + kA^* \,) - (\, A + kA^* \,)A^{\dagger}A \,] = 2r[\, A, \;\; A^* \,] - 2r(A)$.
(e) *The following five statements are equivalent*:
 (1) A *is EP.*
 (2) $AA^{\dagger}(\, A + kA^{\dagger} \,) = (\, A + kA^{\dagger} \,)AA^{\dagger}$.
 (3) $A^{\dagger}A(\, A + kA^{\dagger} \,) = (\, A + kA^{\dagger} \,)A^{\dagger}A$.
 (4) $AA^{\dagger}(\, A + kA^* \,) = (\, A + kA^* \,)AA^{\dagger}$.
 (5) $A^{\dagger}A(\, A + kA^* \,) = (\, A + kA^* \,)A^{\dagger}A$.

Proof. We only show Part (a). Notice that both $AA^{\dagger}$ and $A^{\dagger}A$ are idempotent matrices. Thus by (4.1) we get

$$\begin{aligned} & r[\, AA^{\dagger}(\, A + kA^{\dagger} \,) - (\, A + kA^{\dagger} \,)AA^{\dagger} \,] \\ & = r\begin{bmatrix} AA^{\dagger}(\, A + kA^{\dagger} \,) \\ AA^{\dagger} \end{bmatrix} + r[\, (\, A + kA^{\dagger} \,)AA^{\dagger}, \;\; AA^{\dagger} \,] - r(AA^{\dagger}) - r(A^{\dagger}A) \\ & = r\begin{bmatrix} A + kAA^{\dagger}A^{\dagger} \\ A^* \end{bmatrix} + r[\, A^2A^{\dagger} + kA^{\dagger}, \;\; A \,] - 2r(A) \\ & = r\begin{bmatrix} A \\ A^* \end{bmatrix} + r[\, A^*, \;\; A \,] - 2r(A) = 2r[\, A, \;\; A^* \,] - 2r(A), \end{aligned}$$

establishing Part (a). $\square$

When $k = 1$, the corresponding result in Theorem 6.2(e) was established by Campbell and Mayer [20].

Theorem 6.3. *Let $A \in \mathcal{C}^{m \times m}$ and $0 \neq k \in \mathcal{C}$ be given. Then*
(a) $r[\, AA^{\dagger}(\, AA^* + kA^*A \,) - (\, AA^* + kA^*A \,)A^{\dagger}A \,] = 2r[\, A, \;\; A^* \,] - 2r(A)$.
(b) $AA^{\dagger}(\, AA^* + kA^*A \,) = (\, AA^* + kA^*A \,)A^{\dagger}A \Leftrightarrow A$ *is EP.*

Proof. Follows from (4.1) by noting that both $AA^{\dagger}$ and $A^{\dagger}A$ are idempotent matrices. $\square$

In an earlier paper by Meyer [93] and a recent paper by Hartwig and Katz [61], they established a necessary and sufficient condition for a block triangular matrix to be EP. Their work now can be extended to the following general settings.

Corollary 6.4. *Let $A \in \mathcal{C}^{m \times m}$, $B \in \mathcal{C}^{m \times k}$, and $D \in \mathcal{C}^{k \times k}$ be given, and let $M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$. Then*

$$r(\, MM^{\dagger} - M^{\dagger}M \,) = 2r\begin{bmatrix} A & A^* & B & 0 \\ 0 & B^* & D & D^* \end{bmatrix} - 2r(M). \tag{6.3}$$

In particular,
(a) *If both A and D are EP, then*

$$r(\, MM^{\dagger} - M^{\dagger}M \,) = 2r[\, A, \;\; B \,] + 2r\begin{bmatrix} B \\ D \end{bmatrix} - 2r\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}.$$

(b) *If $R(B) \subseteq R(A)$ and $R(B^*) \subseteq R(D^*)$, then*

$$r(\, MM^{\dagger} - M^{\dagger}M \,) = 2r[\, A, \;\; A^* \,] + 2r[\, D, \;\; D^* \,] - 2r(A) - 2r(D).$$

(c) (Meyer [93], Hartwig and Katz [61]) *M is EP if and only if both A and D are EP, and $R(B) \subseteq R(A)$ and $R(B^*) \subseteq R(D^*)$. In that case, $MM^{\dagger} = \begin{bmatrix} AA^{\dagger} & 0 \\ 0 & DD^{\dagger} \end{bmatrix}$.*

Proof. Follows immediately from Theorem 6.1 by putting M in it. $\square$

Corollary 6.5. *Let*

$$M = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ & A_{22} & \cdots & A_{2k} \\ & & \ddots & \vdots \\ & & & A_{kk} \end{bmatrix} \in \mathcal{C}^{m \times n}, \qquad A_{ij} \in \mathcal{C}^{m_i \times m_j}$$

be given. Then M is EP if and only if $A_{11},\ A_{22}\cdots,\ A_{nn}$ are EP, and

$$R(A_{ij}) \subseteq R(A_{ii}), \quad \text{and} \quad R(A_{ij}^*) \subseteq R(A_{jj}^*), \qquad i,\ j = 1,\ \cdots,\ n.$$

In that case, $MM^\dagger = \mathrm{diag}(\, A_{11}A_{11}^\dagger,\ A_{22}A_{22}^\dagger,\ \cdots,\ A_{nn}A_{nn}^\dagger\,)$.

Proof. Follows from Theorem 6.1(a) by putting M in it. □

We leave the verification of the following result to the reader. Let $A,\ B \in \mathcal{C}^{m\times m}$ be given, and let $M = \begin{bmatrix} A & B \\ B & A \end{bmatrix}$. Then

$$\begin{aligned} & r(\, MM^\dagger - M^\dagger M\,) \\ & = 2r[\, A+B,\ (A+B)^*\,] + 2r[\, A-B,\ (A-B)^*\,] - 2r(\, A+B\,) - 2r(\, A-B\,) \\ & = r[\,(\, A+B\,)(\, A+B\,)^\dagger - (\, A+B\,)^\dagger(\, A+B\,)\,] + r[\,(\, A-B\,)(\, A-B\,)^\dagger - (\, A-B\,)^\dagger(\, A-B\,)\,]. \end{aligned}$$

In particular, M is EP if and only if $A \pm B$ are EP.

A parallel concept to EP matrices is so-called conjugate EP matrices. A matrix A is said to be conjugate EP if $R(A) = R(A^T)$. If matrices considered are real, then EP matrices and conjugate EP matrices are identical. Much similar to EP matrices, conjugate EP matrices also have some nice properties. One of the basic and nice properties related to a conjugate EP matrix A is $AA^\dagger = \overline{A^\dagger A}$ (see the series work [84], [85], [86], [87], and [88] by Meenakshi and Indira). This equality motivates us to find the following results.

Theorem 6.6. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then*

$$r(\, AA^\dagger - \overline{A^\dagger A}\,) = 2r[\, A,\ A^T\,] - 2r(A). \tag{6.4}$$

In particular,

(a) $AA^\dagger = \overline{A^\dagger A} \Leftrightarrow r[\, A,\ A^T\,] = r(A) \Leftrightarrow R(A) = R(A^T)$, *i.e., A is conjugate EP.*

(b) $AA^\dagger - \overline{A^\dagger A}$ *is nonsingular* $\Leftrightarrow r[\, A,\ A^T\,] = 2r(A) = m \Leftrightarrow R(A) \oplus R(A^T) = \mathcal{C}^m$.

Proof. Since $A^\dagger A$ and $\overline{A^\dagger A}$ are idempotent, applying (3.1) to $AA^\dagger - \overline{A^\dagger A}$, we obtain

$$\begin{aligned} r(\, AA^\dagger - \overline{A^\dagger A}\,) &= r\begin{bmatrix} AA^\dagger \\ \overline{A^\dagger A} \end{bmatrix} + r[\, AA^\dagger,\ \overline{A^\dagger A}\,] - r(AA^\dagger) - r(\overline{A^\dagger A}) \\ &= r\begin{bmatrix} A^\dagger \\ \overline{A} \end{bmatrix} + r[\, A,\ \overline{A^\dagger}\,] - 2r(A) \\ &= r\begin{bmatrix} A^* \\ \overline{A} \end{bmatrix} + r[\, \overline{A},\ A^*\,] - 2r(A) = 2r[\, A,\ A^T\,] - 2r(A), \end{aligned}$$

which is exactly (6.4). The results in Parts (a) and (b) follow immediately from (6.4). □

Corollary 6.7. *Let $A \in \mathcal{C}^{m\times m}$, $B \in \mathcal{C}^{m\times k}$, and $D \in \mathcal{C}^{k\times k}$ be given, and denote $M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$. Then*

$$r(\, MM^\dagger - \overline{M^\dagger M}\,) = 2r\begin{bmatrix} A & A^T & B & 0 \\ 0 & B^T & D & D^T \end{bmatrix} - 2r(M).$$

In particular, M is conjugate EP if and only if A and D are con-EP, and $R(B) \subseteq R(A)$ and $R(B^T) \subseteq R(D^T)$.

Proof. Follows from Theorem 6.6 by putting M in it. □

The work in Theorem 6.1 can be extended to matrix expressions that involve powers of a matrix.

Theorem 6.8. *Let $A \in \mathcal{C}^{m\times m}$ be given and k be an integer with $k \geq 2$. Then*

$$r(\, A^kA^\dagger - A^\dagger A^k\,) = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} + r[\, A^k,\ A^*\,] - 2r(A). \tag{6.5}$$

In particular,

(a) $r(\, A^kA^\dagger - A^\dagger A^k \,) = r[\, A, \ A^* \,] - 2r(A)$, *if* $r(A) = r(A^2)$.

(b) $A^kA^\dagger = A^\dagger A^k \Leftrightarrow r\begin{bmatrix} A^k \\ A^* \end{bmatrix} = r[\, A^k, \ A^* \,] = r(A) \Leftrightarrow R(A^k) \subseteq R(A^*)$ *and* $R[(A^k)^*] \subseteq R(A)$.

(c) $A^kA^\dagger - A^\dagger A^k$ *is nonsingular* $\Leftrightarrow r\begin{bmatrix} A^k \\ A^* \end{bmatrix} = r[\, A^k, \ A^* \,] = 2r(A) = m \Leftrightarrow r(A^k) = r(A)$ *and* $R(A) \oplus R(A^*) = \mathcal{C}^m$.

(d) $r(A) = r(A^2)$ *and* $A^kA^\dagger = A^\dagger A^k \Leftrightarrow A$ *is EP.*

Proof. Writing $A^kA^\dagger - A^\dagger A^k = -[\, (A^\dagger A)A^{k-1} - A^{k-1}(AA^\dagger) \,]$ and applying Eq. (4.1) to it, we obtain

$$\begin{aligned} r(\, A^kA^\dagger - A^\dagger A^k \,) &= r\begin{bmatrix} A^\dagger A^k \\ AA^\dagger \end{bmatrix} + r[\, A^kA^\dagger, \ A^\dagger A \,] - r(AA^\dagger) - r(A^\dagger A) \\ &= r\begin{bmatrix} A^k \\ A^\dagger \end{bmatrix} + r[\, A^k, \ A^\dagger \,] - 2r(A) = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} + r[\, A^k, \ A^* \,] - 2r(A), \end{aligned}$$

as required for (6.5). The results in Parts (a)—(d) follow immediately from (6.5). □

¿From the result in Theorem 6.8(a) we can extend the concept of EP matrix to power case: A square matrix A is said to be *k-power-EP* if $R(A^k) \subseteq R(A^*)$ and $R[(A^k)^*] \subseteq R(A)$, where $k \geq 2$. It is believed that power-EP matrices, as a special type of matrices, might also have some more interesting properties. But we do not intend further to discuss power-EP matrices and the related topics in this monograph. As an exercise, we leave the verification of the following result to the reader.

As an application, now we let $M = \begin{bmatrix} A & B \\ B & A \end{bmatrix}$, where $A, \ B \in \mathcal{C}^{m \times m}$. Then

$$\begin{aligned} & r(\, M^kM^\dagger - M^\dagger M^k \,) \\ & = 2r[\, (A+B)^k, \ (A+B)^* \,] + 2r[\, (A-B)^k, \ (A-B)^* \,] - 2r(\, A+B \,) - 2r(\, A-B \,) \\ & = r[\, (\, A+B \,)^k(\, A+B \,)^\dagger - (\, A+B \,)^\dagger(\, A+B \,) \,] + r[\, (\, A-B \,)(\, A-B \,)^\dagger - (\, A-B \,)^\dagger(\, A-B \,) \,]. \end{aligned}$$

In particular, M is k-power-EP if and only if $A \pm B$ are k-power-EP.

In general for any square matrix A and a polynomial $p(x)$, there is

$$r[\, p(A)A^\dagger - A^\dagger p(A) \,] = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} + r[\, A^k, \ A^* \,] - 2r(A), \tag{6.6}$$

and $p(A)A^\dagger = A^\dagger p(A)$ holds if and only if $R[p(A)] \subseteq R(A^*)$ and $R[p(A)^*] \subseteq R(A)$.

Theorem 6.9. *Let* $A \in \mathcal{C}^{m \times m}$ *be given and* k *be an integer with* $k \geq 2$. *Then*

(a) $r[\, A(A^k)^\dagger - (A^k)^\dagger A \,] = r\begin{bmatrix} A^k \\ A^kA^* \end{bmatrix} + r[\, A^k, \ A^*A^k \,] - 2r(A^k)$.

(b) $r[\, A(A^k)^\dagger - (A^k)^\dagger A \,] = 2r[\, A, \ A^* \,] - 2r(A)$, *if* $r(A) = r(A^2)$.

(c) $A(A^k)^\dagger = (A^k)^\dagger A \Leftrightarrow r\begin{bmatrix} A^k \\ A^kA^* \end{bmatrix} = r[\, A^k, \ A^*A^k \,] = r(A^k) \Leftrightarrow R(A^k) = R(A^*A^k)$ *and* $R[(A^k)^*] = R[A(A^k)^*]$.

(d) *If* $A(A^k)^\dagger = (A^k)^\dagger A$, *then* $A^kA^\dagger = A^\dagger A^k$.

(e) $r(A) = r(A^2)$ *and* $A(A^k)^\dagger = (A^k)^\dagger A \Leftrightarrow A$ *is EP.*

Proof. It follows by (2.2) and block elementary operations that

$$\begin{aligned} r[\, A(A^k)^\dagger - (A^k)^\dagger A \,] &= r\begin{bmatrix} (A^k)^*A^k(A^k)^* & 0 & (A^k)^* \\ 0 & -(A^k)^*A^k(A^k)^* & (A^k)^*A \\ A(A^k)^* & (A^k)^* & 0 \end{bmatrix} - 2r(A^k) \\ &= r\begin{bmatrix} (A^k)^*A^k(A^k)^* & (A^k)^*A^{k-1}(A^k)^* & (A^k)^* \\ 0 & 0 & (A^k)^*A \\ A(A^k)^* & (A^k)^* & 0 \end{bmatrix} - 2r(A^k) \\ &= r\begin{bmatrix} A^k \\ A^kA^* \end{bmatrix} + r[\, A^k, \ A^*A^k \,] - 2r(A^k), \end{aligned}$$

establishing Part (a). Parts (b), (c) and (e) follow immediately from Part (a). Combining Part (c) and Theorem 6.8(a) yields the implication in Part (d). □

Theorem 6.10. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then*

(a) $r[\, A(\, AA^\dagger - A^\dagger A\,) - (\, AA^\dagger - A^\dagger A\,)A\,] = r\begin{bmatrix} A & A^2 & A^* \\ A^2 & 0 & 0 \\ A^* & 0 & 0 \end{bmatrix} - 2r(A).$

(b) *A commutes with $AA^\dagger - A^\dagger A \Leftrightarrow r\begin{bmatrix} A & A^2 & A^* \\ A^2 & 0 & 0 \\ A^* & 0 & 0 \end{bmatrix} = 2r(A).$*

(c) $r[\, A(\, AA^\dagger - A^\dagger A\,) - (\, AA^\dagger - A^\dagger A\,)A\,] = 2r[\, A,\ A^*\,] - 2r(A),$ *if* $r(A) = r(A^2)$.

(d) *$r(A) = r(A^2)$ and A commutes with $AA^\dagger - A^\dagger A \Leftrightarrow A$ is EP.*

Proof. Notice that $A(\, AA^\dagger - A^\dagger A\,) - (\, AA^\dagger - A^\dagger A\,)A = A^2A^\dagger + A^\dagger A^2 - 2A$. Thus according to (2.2), we get

$$\begin{aligned}
r(\, A^2A^\dagger + A^\dagger A^2 - 2A\,) &= r\begin{bmatrix} A^*AA^* & 0 & A^* \\ 0 & A^*AA^* & A^*A^2 \\ A^2A^* & A^* & 2A \end{bmatrix} - 2r(A) \\
&= r\begin{bmatrix} A^*A & 0 & A^* \\ 0 & AA^* & A^2 \\ A^2 & A^* & 2A \end{bmatrix} - 2r(A) \\
&= r\begin{bmatrix} 0 & 0 & A^* \\ -A^3 & AA^* & A^2 \\ -A^2 & A^* & 2A \end{bmatrix} - 2r(A) \\
&= r\begin{bmatrix} 0 & 0 & A^* \\ 0 & 0 & -A^2 \\ -A^2 & A^* & 2A \end{bmatrix} - 2r(A) \\
&= r\begin{bmatrix} A & A^2 & A^* \\ A^2 & 0 & 0 \\ A^* & 0 & 0 \end{bmatrix} - 2r(A),
\end{aligned}$$

establishing Part (a). Parts (b), (c) and (d) are are direct consequences of Part (a). □

Theorem 6.11. *Let $A \in \mathcal{C}^{m\times m}$ be given and k be an integer with $k \geq 2$. Then*

(a) $r[\,(AA^\dagger)(A^*A) - (A^*A)(AA^\dagger)\,] = 2r[\, A,\ A^*A^2\,] - 2r(A).$

(b) $r[\,(A^\dagger A)(AA^*) - (AA^*)(A^\dagger A)\,] = 2r\begin{bmatrix} A \\ A^2A^* \end{bmatrix} - 2r(A).$

(c) $r[\,(AA^\dagger)(A^*A)^k - (A^*A)^k(AA^\dagger)\,] = 2r[\, A,\ (A^*A)^kA\,] - 2r(A).$

(d) $r[\,(A^\dagger A)(AA^*)^k - (AA^*)^k(A^\dagger A)\,] = 2r\begin{bmatrix} A \\ A(AA^*)^k \end{bmatrix} - 2r(A).$

(e) *$AA^\dagger$ commutes with $A^*A \Leftrightarrow R(A^*A^2) \subseteq R(A)$.*

(f) *$A^\dagger A$ commutes with $AA^* \Leftrightarrow R[A(A^2)^*] \subseteq R(A^*)$.*

(g) $r[\,(AA^\dagger)(A^*A) - (A^*A)(AA^\dagger)\,] = r[\,(A^\dagger A)(AA^*) - (AA^*)(A^\dagger A)\,] = 2r[\, A,\ A^*\,] - 2r(A),$ *if* $r(A) = r(A^2)$.

(h) *$r(A) = r(A^2)$ and $AA^\dagger$ commutes with $A^*A \Leftrightarrow r(A) = r(A^2)$ and $A^\dagger A$ commutes with $AA^* \Leftrightarrow A$ is EP.*

Proof. Notice that both $AA^\dagger$ and $A^\dagger A$ are idempotent. The two rank equalities in Parts (a)—(d) can trivially be derived from (4.1). The results in Parts (e)—(h) are direct consequences of Parts (a) and (b). □

Theorem 6.12. *Let $A \in \mathcal{C}^{m\times m}$ given and k be an integer with $k \geq 2$. Then*

(a) $r(\, A^\dagger A^*A - A^*AA^\dagger\,) = r[\, A(A^*A) - (A^*A)A\,].$

(b) $r(\, A^\dagger AA^* - AA^*A^\dagger\,) = r[\, A(AA^*) - (AA^*)A\,].$

(c) $r[\, A^\dagger(A^*A)^k - (A^*A)^kA^\dagger\,] = r[\, A(A^*A)^k - (A^*A)^kA\,].$

(d) $r[\,A^{\dagger}(AA^*)^k - (AA^*)^kA^{\dagger}\,] = r[\,A(AA^*)^k - (AA^*)^kA\,]$.

(e) *The following statements are equivalent*:

(1) $A^{\dagger}$ *commutes with* AA^*.

(2) $A^{\dagger}$ *commutes with* A^*A.

(3) $A^{\dagger}$ *commutes with* $(AA^*)^k$.

(4) $A^{\dagger}$ *commutes with* $(A^*A)^k$.

(5)[42] A *commutes with* AA^*.

(6)[42] A *commutes with* A^*A.

(7)[42] A *commutes with* $(AA^*)^k$.

(8)[42] A *commutes with* $(A^*A)^k$.

(9)[42] A *is normal, i.e.,* $AA^* = A^*A$.

Proof. Notice that $A^{\dagger}A^*A - A^*AA^{\dagger} = -(\,A^* - A^{\dagger}A^*A\,)$. Thus by (2.1) we find that

$$\begin{aligned} r(\,A^* - A^{\dagger}A^*A\,) &= r\begin{bmatrix} A^*AA^* & A^*A^*A \\ A^* & A^* \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} 0 & A^*A^*A - A^*AA^* \\ A^* & 0 \end{bmatrix} - r(A) = r(\,AA^*A - A^*A^2\,), \end{aligned}$$

establishing Part (a). Similarly we can establish Parts (b)—(c). The equivalence of (1) and (5), (2) and (6), (3) and (7), (4) and (8) in Part (c) follow from the four formulas in Parts (a)—(b). The equivalence of (5)—(9) in Part (c) were presented in [42]. □

A square matrix A is said to be bi-EP, if A and its Moore-Penrose inverse $A^{\dagger}$ satisfy $(AA^{\dagger})(A^{\dagger}A) = (A^{\dagger}A)(AA^{\dagger})$. This special type of matrices were examined by Campbell and Meyer [20], Hartwig and Spindelböck [63], [64].

Just as for EP matrices and conjugate EP matrices, bi-EP matrices can also be characterized by a rank equality.

Theorem 6.13. *Let* $A \in \mathcal{C}^{m\times m}$ *be given. Then*

(a) $r[\,(AA^{\dagger})(A^{\dagger}A) - (A^{\dagger}A)(AA^{\dagger})\,] = 2r[\,A,\ A^*\,] + 2r(A^2) - 4r(A)$.

(b) $r[\,A^2 - A^2(A^{\dagger})^2A^2\,] = r[\,A,\ A^*\,] + r(A^2) - 2r(A)$.

(c) $r[\,(AA^{\dagger})(A^{\dagger}A) - (A^{\dagger}A)(AA^{\dagger})\,] = 2r[\,A,\ A^*\,] - 2r(A)$, *if* $r(A) = r(A^2)$.

(d) $r[\,A^2 - A^2(A^{\dagger})^2A^2\,] = r[\,A,\ A^*\,] - r(A)$, *if* $r(A) = r(A^2)$.

(e) *The following four statements are equivalent*:

(1) $(AA^{\dagger})(A^{\dagger}A) = (A^{\dagger}A)(AA^{\dagger})$, *i.e.,* A *is bi-EP.*

(2) $(A^{\dagger})^2 \in \{(A^2)^-\}$.

(3) $r[\,A,\ A^*\,] = 2r(A) - r(A^2)$.

(4) $\dim[R(A) \cap R(A^*)] = r(A^2)$.

(f)[64] A *is bi-EP and* $r(A) = r(A^2) \Leftrightarrow A^2(A^{\dagger})^2A^2 = A^2$ *and* $r(A) = r(A^2) \Leftrightarrow A$ *is EP.*

Proof. Note that both $AA^{\dagger}$ and $A^{\dagger}A$ are Hermitian idempotent and $R(A^{\dagger}) = R(A^*)$. We find by (3.29) that

$$\begin{aligned} r[\,(AA^{\dagger})(A^{\dagger}A) - (A^{\dagger}A)(AA^{\dagger})\,] &= 2r[\,AA^{\dagger},\ A^{\dagger}A\,] + 2r[\,(AA^{\dagger})(A^{\dagger}A)\,] - 2r(AA^{\dagger}) - 2r(A^{\dagger}A) \\ &= 2r[\,A,\ A^*\,] + 2r(A^{\dagger}A^{\dagger}) - 4r(A) \\ &= 2r[\,A,\ A^*\,] + 2r(A^2) - 4r(A), \end{aligned}$$

establishing Part (a). Applying (2.8) and then the rank cancellation laws in (1.8) to $A^2 - A^2(A^{\dagger})^2A^2$, we obtain

$$\begin{aligned} r[\,A^2 - A^2(A^{\dagger})^2A^2\,] &= r\begin{bmatrix} A^*A^* & A^*AA^* & 0 \\ A^*AA^* & 0 & A^*A^2 \\ 0 & A^2A^* & -A^2 \end{bmatrix} - 2r(A) \\ &= r\begin{bmatrix} A^*A^* & A^*AA^* & 0 \\ A^*AA^* & A^*A^2A^* & 0 \\ 0 & 0 & -A^2 \end{bmatrix} - 2r(A) \end{aligned}$$

$$
\begin{aligned}
&= r\begin{bmatrix} A^*A^* & A^*AA^* \\ A^*AA^* & A^*A^2A^* \end{bmatrix} + r(A^2) - 2r(A) \\
&= r\begin{bmatrix} A^*A^* & A^*A \\ AA^* & A^2 \end{bmatrix} + r(A^2) - 2r(A) \\
&= r\left([\, A, \ A^* \,]^*[\, A, \ A^* \,] \right) + r(A^2) - 2r(A) \\
&= r[\, A, \ A^* \,] + r(A^2) - 2r(A),
\end{aligned}
$$

as required for Part (b). The equivalence of (1)—(3) in Part (e) follows from the two formulas in Parts (a) and (b) The equivalence of (3) and (4) in Part (e) follows from a well-known rank formula $r[\, A, \ B\,] = r(A) + r(B) - \dim[\, R(A) \cap R(B)\,]$. □

The above work can also be extended to the conjugate case.

Theorem 6.14. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then*

$$r[\,(AA^\dagger)(\overline{A^\dagger A}) - (\overline{A^\dagger A})(AA^\dagger)\,] = 2r[\, A, \ A^T\,] + 2r(A\overline{A}) - 4r(A).$$

In particular,

$$(AA^\dagger)(\overline{A^\dagger A}) = (\overline{A^\dagger A})(AA^\dagger) \Leftrightarrow r[\, A, \ A^T\,] = 2r(A) - r(A\overline{A}).$$

Proof. Follows from (3.29) by noticing that both $AA^\dagger$ and $\overline{A^\dagger A}$ are idempotent. □.

Based on the above results, a parallel concept to bi-EP matrix now can be introduced: A square matrix A is said to be *conjugate bi-EP* if $(AA^\dagger)(\overline{A^\dagger A}) = (\overline{A^\dagger A})(AA^\dagger)$. The properties and applications of this special type of matrices remain to further study.

We next consider rank equalities related to star-dagger matrices. A square matrix A is said to be star-dagger if $A^*A^\dagger = A^\dagger A^*$. This special type of matrices were well investigated by Hartwig and Spindelböck [64], and later by Meenakshi and Rajian [90].

Theorem 6.15. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then*

(a) $r(\, A^*A^\dagger - A^\dagger A^*\,) = r(\, AA^*A^2 - A^2A^*A\,)$.
(b) $r(\, AA^*A^\dagger A - AA^\dagger A^*A\,) = r(\, AA^*A^2 - A^2A^*A\,)$.
(c) $r(\, A^*A^\dagger - A^\dagger A^*\,) = r(\, AA^* - A^*A\,)$, *if A is EP.*
(d) *The following statements are equivalent (Hartwig and Spindelböck [64]):*
  (1) $A^*A^\dagger = A^\dagger A^*$, *i.e., A is star-dagger.*
  (2) $AA^*A^\dagger A = AA^\dagger A^*A$.
  (3) $AA^*A^2 = A^2A^*A$.
(e)[95] *A is both EP and star-dagger $\Leftrightarrow$ A is normal.*

Proof. We find by (2.2) that

$$
\begin{aligned}
r(\, A^*A^\dagger - A^\dagger A^*\,) &= r\begin{bmatrix} -A^*AA^* & 0 & A^* \\ 0 & A^*AA^* & A^*A^* \\ A^*A^* & A^* & 0 \end{bmatrix} - 2r(A) \\
&= r\begin{bmatrix} 0 & 0 & A^* \\ (A^*)^2AA^* - A^*A(A^*)^2 & 0 & \\ 0 & 0 & A \\ A & 0 & 0 \end{bmatrix} - 2r(A) = r(\, AA^*A^2 - A^2A^*A\,),
\end{aligned}
$$

as required for Part (a). Similarly we can establish Part (b) by (2.2). The formula in in Part (c) is derived from Part (a), and Part (d) is direct consequences of Parts (a) and (b). Part (e) comes from Part (c). □

As pointed out by Hartwig and Spindelböck [64], the class of star-dagger matrices are quite inclusive. Normal matrix, partial isometry (i.e., $A^\dagger = A^*$), idempotent matrix, 2-nilpotent matrix, power Hermitian matrix (i.e., $A^* = A^k$), and so on are all special cases of star-dagger matrices, this assertion can easily be seen from the statement (3) in Theorem 6.15(d).

The results in Theorem 6.15 can be extended to general cases. Below are three of them. Their proofs are much similar to that of Theorem 6.15 and are, therefore, omitted.

Theorem 6.16. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then*
(a) $r(\, A^*AA^*A^\dagger - A^\dagger A^*AA^*\,) = r[\,(AA^*)^2A^2 - A^2(A^*A)^2\,]$.
(b) $r[\,(AA^*)^2A^\dagger A - AA^\dagger(A^*A)^2\,] = r[\,(AA^*)^2A^2 - A^2(A^*A)^2\,]$.
(c) $A^*AA^*A^\dagger = A^\dagger A^*AA^* \Leftrightarrow (AA^*)^2A^\dagger A = AA^\dagger(A^*A)^2 \Leftrightarrow (AA^*)^2A^2 = A^2(A^*A)^2$.
(d) *If $A^*A^\dagger = A^\dagger A^*$, then $A^*AA^*A^\dagger = A^\dagger A^*AA^*$.*

Theorem 6.17. *Let $A \in \mathcal{C}^{m\times m}$ be given and k be an integer with $k \geq 2$. Then*
(a) $r[\,(A^*)^kA^\dagger - A^\dagger(A^*)^k\,] = r(\,AA^*A^{k+1} - A^{k+1}A^*A\,)$.
(b) $r[\,A(A^*)^kA^\dagger A - AA^\dagger(A^*)^kA\,] = r(\,AA^*A^{k+1} - A^{k+1}A^*A\,)$.
(c) $(A^*)^kA^\dagger = A^\dagger(A^*)^k \Leftrightarrow A(A^*)^kA^\dagger A = AA^\dagger(A^*)^kA \Leftrightarrow AA^*A^{k+1} = A^{k+1}A^*A$.
(d) *If $A^{k+1} = A$, or $A^{k+1} = 0$, or $AA^* = A^*A$, or $AA^*A = A$, then $(A^*)^kA^\dagger = A^\dagger(A^*)^k$ holds.*

In general for any square matrix A and a polynomial $p(x)$, there is

$$r[\,p(A^*)A^\dagger - A^\dagger p(A^*)\,] = r[\,AA^*p(A)A - Ap(A)A^*A\,]. \tag{6.7}$$

In particular, $p(A^*)A^\dagger = A^\dagger p(A^*)$ holds if and only if $AA^*p(A)A = Ap(A)A^*A$.

Theorem 6.18. *Let $A \in \mathcal{C}^{m\times m}$ be given and k be an integer with $k \geq 2$. Then*
(a) $r[\,A^*(A^k)^\dagger - (A^k)^\dagger A^*\,] = r[\,A^k(A^k)^*A^{k+1} - A^{k+1}(A^k)^*A^k\,]$.
(b) $A^*(A^k)^\dagger = (A^k)^\dagger A^* \Leftrightarrow A^k(A^k)^*A^{k+1} = A^{k+1}(A^k)^*A^k$.

Next are several results on ranks of matrix expressions involving powers of the Moore-Penrose inverse of a matrix.

Theorem 6.19. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then*
(a) $r[\,I_m \pm A^\dagger\,] = r(\,A^2 \pm AA^*A\,) - r(A) + m$.
(b) $r[\,I_m - (A^\dagger)^2\,] = r(\,A^2 + AA^*A\,) + r(\,A^2 - AA^*A\,) - 2r(A) + m$.
(c) $r[\,I_m \pm A^\dagger\,] = r(\,A \pm AA^*\,) - r(A) + m$, *if A is EP.*
(d) $r[\,I_m - (A^\dagger)^2\,] = r(\,A + AA^*\,) + r(\,A - AA^*\,) - 2r(A) + m$, *if A is EP.*
(e) $r[\,I_m \pm A^\dagger\,] = r(\,A \pm A^2\,) - r(A) + m$, *if A is Hermitian.*
(f) $r[\,I_m - (A^\dagger)^2\,] = r(\,A + A^2\,) + r(\,A - A^2\,) - 2r(A) + m$, *if A is Hermitian.*

Proof. By (2.1) we easily obtain

$$\begin{aligned} r(\,I_m - A^\dagger\,) &= r\begin{bmatrix} A^*AA^* & A^* \\ A^* & I_m \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} A^*AA^* - A^*A^* & 0 \\ 0 & I_m \end{bmatrix} - r(A) = r(\,AA^*A - A^2\,) + m - r(A), \end{aligned}$$

and

$$\begin{aligned} r(\,I_m + A^\dagger\,) &= r\begin{bmatrix} -A^*AA^* & A^* \\ A^* & I_m \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} -A^*AA^* - A^*A^* & 0 \\ 0 & I_m \end{bmatrix} - r(A) = r(\,AA^*A + A^2\,) + m - r(A). \end{aligned}$$

Both of the above are exactly Part (a). Next applying (1.12) to $I_m - (A^\dagger)^2$ we obtain

$$\begin{aligned} r[\,I_m - (A^\dagger)^2\,] &= r(\,I_m + A^\dagger\,) + r(\,I_m - A^\dagger\,) - m \\ &= r(\,A^2 + AA^*A\,) + r(\,A^2 - AA^*A\,) - 2r(A) + m, \end{aligned}$$

establishing Part (b). The results in Parts (c)—(f) follow directly from Parts (a) and (b). □

Theorem 6.20. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then*
(a) $r[\,A^\dagger \pm (A^\dagger)^2\,] = r(\,A^2 \pm AA^*A\,) = r[\,A \pm A(A^\dagger)^2A\,]$.

(b) $r[\, A^\dagger - (A^\dagger)^2 \,] = r(\, A - AA^* \,) = r(\, A - A^*A \,)$, *if A is EP.*

(c) $r[\, A^\dagger \pm (A^\dagger)^2 \,] = r(\, A \pm A^2 \,)$, *if A is Hermitian.*

(d) $r[\, A^\dagger - (A^\dagger)^2 \,] = r(A^\dagger) - r[\,(A^\dagger)^2\,]$, *i.e.*, $(A^\dagger)^2 \leq_{rs} A^\dagger \Leftrightarrow r(\, AA^*A - A^2 \,) = r(A) - r(A^2)$, *i.e.*, $A^2 \leq_{rs} AA^*A$.

(e) $(A^\dagger)^2 = A^\dagger \Leftrightarrow AA^*A = A^2 \Leftrightarrow A = (AA^\dagger)(A^\dagger A) \Leftrightarrow (A^\dagger)^2 \in \{A^-\}$.

(f) $(A^\dagger)^2 = A^\dagger \Leftrightarrow AA^* = A^*A = A$, *if A is EP.*

(g) $(A^\dagger)^2 = A^\dagger \Leftrightarrow A^2 = A$, *if A is Hermitian.*

Proof. It follows first from (1.11) that

$$r[\, A^\dagger - (A^\dagger)^2 \,] = r(\, I_m - A^\dagger \,) + r(A) - m,$$
$$r[\, A^\dagger + (A^\dagger)^2 \,] = r(\, I_m + A^\dagger \,) + r(A) - m.$$

Then we have the first two equalities in Part (a) by Theorem 6.15(a). Note that

$$A[\, A^\dagger \pm (A^\dagger)^2 \,]A = A \pm A(A^\dagger)^2A \quad \text{and} \quad A^\dagger[\, A \pm A(A^\dagger)^2A \,]A^\dagger = A^\dagger \pm (A^\dagger)^2.$$

It follows that

$$r[\, A^\dagger \pm (A^\dagger)^2 \,] = r[\, A \pm A(A^\dagger)^2A \,].$$

Thus we have the second equality in Part (a). Parts (b)—(g) follow from Part (a). □

Theorem 6.21. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then*

(a) $r[\, A^\dagger - (A^\dagger)^3 \,] = r(\, A^2 + AA^*A \,) + r(\, A^2 - AA^*A \,) - r(A)$.

(b) $r[\, A^\dagger - (A^\dagger)^3 \,] = r(\, A + AA^* \,) + r(\, A - AA^* \,) - r(A)$, *if A is EP.*

(c) $r[\, A^\dagger - (A^\dagger)^3 \,] = r(\, A + A^2 \,) + r(\, A - A^2 \,) - r(A) = r(\, A - A^3 \,)$, *if A is Hermitian.*

(d) $(A^\dagger)^3 = A^\dagger \Leftrightarrow r(\, A^2 + AA^*A \,) + r(\, A^2 - AA^*A \,) = r(A) \Leftrightarrow R(\, AA^*A + A^2 \,) \cap R(\, AA^*A - A^2 \,) = \{0\}$ *and* $R[(\, AA^*A + A^2 \,)^*] \cap R[(\, AA^*A - A^2 \,)^*] = \{0\}$.

Proof. Applying the rank equality (1.15) to $A^\dagger - (A^\dagger)^3$, we obtain

$$r[\, A^\dagger - (A^\dagger)^3 \,] = r[\, A^\dagger + (A^\dagger)^2 \,] + r[\, A^\dagger - (A^\dagger)^2 \,] - r(A).$$

Then putting Theorem 6.20(a) in it yields Part (a). The results in Parts (b)—(d) follow all from Part (a). □

Theorem 6.22. *Let $A \in \mathcal{C}^{m\times n}$ be given. Then*

(a) $r(\, A^\dagger - A^* \,) = r(\, A - AA^*A \,)$.

(b) $r(\, A^\dagger - A^*AA^* \,) = r(\, A - AA^*AA^*A \,)$.

(c) $r[\, A^\dagger - (A^k)^* \,] = r(\, A - A^kA^*A \,) = r(\, A - AA^*A^k \,)$.

In particular,

(d) $A^\dagger = A^* \Leftrightarrow AA^*A = A$, *i.e.*, *$A$ is partial isometry.*

(e) $A^\dagger = A^*AA^* \Leftrightarrow AA^*AA^*A = A$.

(f) $A^\dagger = (A^k)^* \Leftrightarrow A^kA^*A = AA^*A^k = A$.

Proof. Follows from (2.1). □

Theorem 6.23. *Let $A \in \mathcal{C}^{m\times m}$ be idempotent. Then*

(a) $r(\, A - A^\dagger \,) = 2r[\, A, \ A^* \,] - 2r(A)$.

(b) $r(\, 2A - AA^\dagger - A^\dagger A \,) = 2r[\, A, \ A^* \,] - 2r(A)$.

(c) $r[\,(AA^\dagger)(AA^*) - (AA^*)(AA^\dagger)\,] = r[\,(A^\dagger A)(A^*A) - (A^*A)(A^\dagger A)\,] = 2r[\, A, \ A^* \,] - 2r(A)$.

(d) $r(\, A^\dagger - AA^\dagger A^\dagger A \,) = 2r[\, A, \ A^* \,] - 2r(A)$.

(e) $r(\, A - AA^\dagger A^\dagger A \,) = r[\, A, \ A^* \,] - r(A)$.

(f) *$A^\dagger$ commutes with A^*.*

(g) *$A^\dagger$ commutes with A^*AA^*.*

(h) $(AA^*)^2A^\dagger A = AA^\dagger A(A^*A)^2$.

In particular,

(i) $A = A^\dagger \Leftrightarrow AA^\dagger + A^\dagger A = 2A \Leftrightarrow (AA^\dagger)(AA^*) = (AA^*)(AA^\dagger) \Leftrightarrow (A^\dagger A)(A^*A) = (A^*A)(A^\dagger A) \Leftrightarrow A^\dagger = AA^\dagger A^\dagger A \Leftrightarrow A = AA^\dagger A^\dagger A \Leftrightarrow A$ *is Hermitian.*

Proof. Note that $A, \ A^\dagger \in A\{2\}$ when A is idempotent. Thus we have by (5.1) that

$$\begin{aligned} r(A - A^\dagger) &= r\begin{bmatrix} A \\ A^\dagger \end{bmatrix} + r[\,A, \ A^\dagger\,] - r(A) - r(A^\dagger) \\ &= r\begin{bmatrix} A \\ A^* \end{bmatrix} + r[\,A, \ A^*\,] - 2r(A) = 2r[\,A, \ A^*\,] - 2r(A), \end{aligned}$$

establishing Part (a). Part (b) follows from Theorem 6.10(c), Part (c) follows from Theorem 6.11(c), and Parts (d) and (e) follow from Theorem 6.13(a) and (b). Parts (f)—(h) follow from Theorems 6.15 and 6.16. Part (i) is a direct consequence of Parts (a)—(e). □

Theorem 6.24. *Let $A \in \mathcal{C}^{m\times m}$ be tripotent, that is, $A^3 = A$. Then*

(a) $r(A - A^\dagger) = 2r[\,A, \ A^*\,] - 2r(A)$.
(b) $r(A^2A^\dagger - A^\dagger A^2) = 2r[\,A, \ A^*\,] - 2r(A)$.
(c) $r[\,A(A^2)^\dagger - (A^2)^\dagger A\,] = 2r[\,A, \ A^*\,] - 2r(A)$.
(d) $r[\,A(AA^\dagger - A^\dagger A) - (AA^\dagger - A^\dagger A)A\,] = 2r[\,A, \ A^*\,] - 2r(A)$.
(e) $r[\,(AA^\dagger)(AA^*) - (AA^*)(AA^\dagger)\,] = 2r[\,A, \ A^*\,] - 2r(A)$.
(f) $r[\,(A^\dagger A)(A^*A) - (A^*A)(A^\dagger A)\,] = 2r[\,A, \ A^*\,] - 2r(A)$.
(g) $r[\,(AA^\dagger)(A^\dagger A) - (A^\dagger A)(AA^\dagger)\,] = 2r[\,A, \ A^*\,] - 2r(A)$.
(h) $r[\,A^2 - A^2(A^\dagger)^2A^2\,] = r[\,A, \ A^*\,] - r(A)$.
(i) $(A^*)^2A^\dagger = A^\dagger(A^*)^2$.
(j) *The following nine statements are equivalent:*
 (1) $A = A^\dagger$.
 (2) $A^2A^\dagger = A^\dagger A^2$.
 (3) $A(A^2)^\dagger = (A^2)^\dagger A$.
 (4) $A(AA^\dagger - A^\dagger A) = (AA^\dagger - A^\dagger A)A$.
 (5) $(AA^\dagger)(AA^*) = (AA^*)(AA^\dagger)$.
 (6) $(A^\dagger A)(A^*A) = (A^*A)(A^\dagger A)$.
 (7) $(AA^\dagger)(A^\dagger A) = (A^\dagger A)(AA^\dagger)$.
 (8) $A^2 = A^2(A^\dagger)^2A^2$.
 (9) $R(A) = R(A^*)$, *i.e., A is EP.*

Proof. Note that $A, \ A^\dagger \in A\{2\}$ when A is tripotent. Thus we have by (5.1) that

$$\begin{aligned} r(A - A^\dagger) &= r\begin{bmatrix} A \\ A^\dagger \end{bmatrix} + r[\,A, \ A^\dagger\,] - r(A) - r(A^\dagger) \\ &= r\begin{bmatrix} A \\ A^* \end{bmatrix} + r[\,A, \ A^*\,] - 2r(A) = 2r[\,A, \ A^*\,] - 2r(A), \end{aligned}$$

establishing Part (a). Parts (b)—(h) follow respectively from (6.5), Theorem 6.9(b), Theorem 6.10(c), Theorem 6.11(e), and Theorem 6.13(a) and (b). Part (i) follows from Theorem 6.17(d). Part (j) is a direct consequence of Parts (a)—(h). □

The following result is motivated by a problem of Rao and Mitra [118] on the nonsingularity of a matrix of the form $I + A - A^\dagger A$.

Theorem 6.25. *Let $A \in \mathcal{C}^{m\times m}$ and $1 \neq \lambda \in \mathcal{C}$ be given. Then*

(a) $r(I_m + A - A^\dagger A) = r(I_m + A - AA^\dagger) = r(A^2) - r(A) + m$.
(b) $r(I_m - A - A^\dagger A) = r(I_m - A - AA^\dagger) = r(A^2) - r(A) + m$.
(c) $r(\lambda I_m + A - A^\dagger A) = r(\lambda I_m + A - AA^\dagger) = r[\,(\lambda - 1)I_m + A\,]$.
(d) $r(\mu I_m - A) = r(\mu I_m + I_m - A - A^\dagger A) = r(\mu I_m + I_m - A - AA^\dagger)$, *when* $\mu \neq 0$.
(e)[118] $I_m + A - A^\dagger A$ *is nonsingular* $\Leftrightarrow I_m + A - AA^\dagger$ *is nonsingular* $\Leftrightarrow r(A^2) = r(A)$.

Proof. Applying (2.1) and then (1.8), we find that

$$\begin{aligned} r(\, I_m + A - A^{\dagger}A\,) &= r\begin{bmatrix} A^*AA^* & A^*A \\ A^* & I_m + A \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} AA^* & A \\ A^* & I_m + A \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} 0 & A \\ -AA^* & I_m \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} A^2A^* & 0 \\ 0 & I_m \end{bmatrix} - r(A) = m + r(A^2) - r(A). \end{aligned}$$

By symmetry, we also get $r(\, I_m + A - AA^{\dagger}\,) = m + r(A^2) - r(A)$. Both of them are the result in Part (a). Replace A by $-A$ to yield Part (b). Part (c) is derived by (2.1) and (1.11). Replace λ by $\mu + 1$ in Part (c) to yield Part (d). Part (e) is a direct consequence of Part (a). $\square$

The result in Theorem 6.26(d) reveals an interesting fact that a square matrix A has the same nonzero eigenvalues as the matrix $A + A^{\dagger}A - I_m$ (or $A + AA^{\dagger} - I_m$) has. Of course, this result is trivial when A is nonsingular.

Theorem 6.26. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then*

(a) $r(\, I_m + A^k - A^{\dagger}A\,) = r(\, I_m + A^k - AA^{\dagger}\,) = r(A^{k+1}) - r(A) + m.$

(b) $r(\, I_m - A^k - A^{\dagger}A\,) = r(\, I_m - A^k - AA^{\dagger}\,) = r(A^{k+1}) - r(A) + m.$

(c) $I_m + A^k - A^{\dagger}A$ *is nonsingular* $\Leftrightarrow I_m + A^k - AA^{\dagger}$ *is nonsingular* $\Leftrightarrow I_m - A^k - A^{\dagger}A$ *is nonsingular* $\Leftrightarrow I_m - A^k - AA^{\dagger}$ *is nonsingular* $\Leftrightarrow r(A^{k+1}) = r(A)$, *i.e.*, $r(A) = r(A^2)$.

(d) $r(\, I_m + A^k - A^{\dagger}A\,) = m - r(A) \Leftrightarrow A^{k+1} = 0.$

Proof. We only show the first equality in Part (a). Applying (2.1) and then (1.8), we find that

$$\begin{aligned} r(\, I_m + A^k - A^{\dagger}A\,) &= r\begin{bmatrix} A^*AA^* & A^*A \\ A^* & I_m + A^k \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} AA^* & A \\ A^* & I_m + A^k \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} 0 & A \\ -A^kA^* & I_m \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} A^{k+1}A^* & 0 \\ 0 & I_m \end{bmatrix} - r(A) = m + r(A^{k+1}) - r(A). \qquad \square \end{aligned}$$

The rank equalities in Theorem 6.26 are still valid when replacing the Moore-Penrose inverse of A by any inner inverse A^- of A. We shall prove this in Chapter 23.

Theorem 6.27. *Let $A \in \mathcal{C}^{m\times m}$ be given. Then*

(a) $r(\, I_m - AA^{\dagger} - A^{\dagger}A\,) = 2r(A^2) - 2r(A) + m.$

(b) $I_m - AA^{\dagger} - A^{\dagger}A$ *is nonsingular* $\Leftrightarrow r(A^2) = r(A).$

(c) $AA^{\dagger} + A^{\dagger}A = I_m \Leftrightarrow m$ *is even*, $r(A) = m/2$ *and* $A^2 = 0.$

Proof. Apply (3.8) to $I_m - AA^{\dagger} - A^{\dagger}A$ to yield

$$\begin{aligned} r(\, I_m - AA^{\dagger} - A^{\dagger}A\,) &= r(AA^{\dagger}A^{\dagger}A) + r(A^{\dagger}AAA^{\dagger}) - r(AA^{\dagger}) - r(A^{\dagger}A) + m \\ &= 2r(A^2) - 2r(A) + m, \end{aligned}$$

as required in Part (a). Part (b) is obvious from Part (a). According to the rank formula in Part (a), the equality $AA^{\dagger} + A^{\dagger}A = I_m$ holds if and only if $2r(A^2) - 2r(A) + m = 0$. This rank equality implies that m must be even and $r(A^2) = r(A) - m/2$. Contracting this rank equality with the Frobenius rank inequality $r(A^2) \geq 2r(A) - m$, we get Part (c). $\square$

Chapter 7

Rank equalities for matrices and their Moore-Penrose inverses

We consider in this chapter ranks of various matrix expressions that involve two or more matrices and their Moore-Penrose inverses, and present their various consequences, which can reveal a series of intrinsic properties related to Moore-Penrose inverses of matrices. Most of the results obtained in this chapter are new and are not considered before.

Theorem 7.1. *Let* $A,\ B \in \mathcal{C}^{m\times m}$ *be given. Then*

(a) *The rank of* $AA^{\dagger}B - BA^{\dagger}A$ *satisfies*

$$r(\,AA^{\dagger}B - BA^{\dagger}A\,) = r\begin{bmatrix} A \\ A^*B \end{bmatrix} + r[\,A,\ BA^*\,] - 2r(A). \tag{7.1}$$

(b) $AA^{\dagger}B = BA^{\dagger}A \Leftrightarrow r\begin{bmatrix} A \\ A^*B \end{bmatrix} = r[\,A,\ BA^*\,] = r(A) \Leftrightarrow R(BA^*) \subseteq R(A)$ *and* $R(B^*A) \subseteq R(A^*)$.

(c) $AA^{\dagger}B - BA^{\dagger}A$ *is nonsingular* $\Leftrightarrow r\begin{bmatrix} A \\ A^*B \end{bmatrix} = r[\,A,\ BA^*\,] = 2r(A) = m \Leftrightarrow R(A) \oplus R(BA^*) = \mathcal{C}^m$ *and* $R(AB^*) = R(A) \Leftrightarrow R(A^*) \oplus R(B^*A) = \mathcal{C}^m$ *and* $R(A^*B) = R(A^*)$.

Proof. Note that $AA^{\dagger}$ and $A^{\dagger}A$ are idempotent and $R(A^{\dagger}) = R(A^*)$. We have by Eq. (4.1) that

$$\begin{aligned} r(\,AA^{\dagger}B - BA^{\dagger}A\,) &= r\begin{bmatrix} AA^{\dagger}B \\ A^{\dagger}A \end{bmatrix} + r[\,BA^{\dagger}A,\ AA^{\dagger}\,] - r(AA^{\dagger}) - r(A^{\dagger}A) \\ &= r\begin{bmatrix} A^{\dagger}B \\ A \end{bmatrix} + r[\,BA^{\dagger},\ A\,] - 2r(A) \\ &= r\begin{bmatrix} A^*B \\ A \end{bmatrix} + r[\,BA^*,\ A\,] - 2r(A), \end{aligned}$$

establishing (7.1). The results in Parts (b) and (c) follow from it. $\square$

Clearly the results in Theorems 6.1 and 6.7 are special cases of the above theorem.

Theorem 7.2. *Let* $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ *and* $C \in \mathcal{C}^{l\times n}$ *be given. Then*

(a) $r(\,AA^{\dagger} - BB^{\dagger}\,) = 2r[\,A,\ B\,] - r(A) - r(B)$.

(b) $r(\,A^{\dagger}A - C^{\dagger}C\,) = 2r\begin{bmatrix} A \\ C \end{bmatrix} - r(A) - r(C)$.

(c) $r(\,AA^{\dagger} + BB^{\dagger}\,) = r[\,A,\ B\,]$, *that is,* $R(\,AA^{\dagger} + BB^{\dagger}\,) = R[\,A,\ B\,]$.

(d) $r(\,A^{\dagger}A + C^{\dagger}C\,) = r\begin{bmatrix} A \\ C \end{bmatrix}$, *that is,* $R(\,A^{\dagger}A + C^{\dagger}C\,) = R[\,A^*,\ C^*\,]$.

In particular,

(e) $AA^{\dagger} = BB^{\dagger} \Leftrightarrow R(A) = R(B)$.

(f) $A^{\dagger}A = C^{\dagger}C \Leftrightarrow R(A^*) = R(C^*)$.
(g) $r(\, AA^{\dagger} - BB^{\dagger}\,) = r(AA^{\dagger}) - r(BB^{\dagger}) \Leftrightarrow R(B) \subseteq R(A)$.
(h) $r(\, A^{\dagger}A - C^{\dagger}C\,) = r(A^{\dagger}A) - r(C^{\dagger}C) \Leftrightarrow R(C^*) \subseteq R(A^*)$.
(i) $r(\, AA^{\dagger} - BB^{\dagger}\,) = m \Leftrightarrow r[\, A, \ B\,] = r(A) + r(B) = m \Leftrightarrow R(A) \oplus R(B) = \mathcal{C}^m$.
(j) $r(\, A^{\dagger}A - C^{\dagger}C\,) = n \Leftrightarrow r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r(C) = n \Leftrightarrow R(A^*) \oplus R(C^*) = \mathcal{C}^n$.

Proof. Note that $AA^{\dagger}$, $A^{\dagger}A$, $BB^{\dagger}$, and $C^{\dagger}C$ are all idempotent. Thus we can easily derive by (3.1) and (3.12) the four rank equalities in Parts (a)—(d). The results in Parts (e)—(j) are direct consequences of Parts (a) and (b). $\square$

Theorem 7.3. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{k\times m}$ be given. Then*

$$r(\, AA^{\dagger}B^{\dagger}B - B^{\dagger}BAA^{\dagger}\,) = 2r[\, A, \ B^*\,] + 2r(BA) - 2r(A) - 2r(B). \tag{7.2}$$

In particular

$$(AA^{\dagger})(B^{\dagger}B) = (B^{\dagger}B)(AA^{\dagger}) \Leftrightarrow r[\, A, \ B^*\,] = r(A) + r(B) - r(BA) \Leftrightarrow \dim[R(A) \cap R(B^*)] = r(BA). \tag{7.3}$$

Proof. Note that $AA^{\dagger}$, $A^{\dagger}A$, $BB^{\dagger}$, and $B^{\dagger}B$ are Hermitian idempotent. Thus we find by (3.29) that

$$\begin{aligned} r[\,(AA^{\dagger})(B^{\dagger}B) - (B^{\dagger}B)(AA^{\dagger})\,] &= 2r[\, AA^{\dagger}, \ B^{\dagger}B\,] + 2r[\,(AA^{\dagger})(B^{\dagger}B)\,] - 2r(AA^{\dagger}) - 2r(B^{\dagger}B) \\ &= 2r[\, A, \ B^*\,] + 2r(BA) - 2r(A) - 2r(B), \end{aligned}$$

as required for (7.2). The result in (7.3) is a direct consequence of (7.2). $\square$

Replace B by B^* in (7.2) to yield an alternative formula

$$r(\, AA^{\dagger}BB^{\dagger} - BB^{\dagger}AA^{\dagger}\,) = 2r[\, A, \ B\,] + 2r(B^*A) - 2r(A) - 2r(B). \tag{7.4}$$

Some interesting consequences can be derived from (7.2) and (7.4). For example, let $B = I_m - A$ in (7.2). Then we get by (1.11)

$$\begin{aligned} & r[\, AA^{\dagger}(I_m - A)^{\dagger}(I_m - A) - (I_m - A)^{\dagger}(I_m - A)AA^{\dagger}\,] \\ &= 2r[\, A, \ I_m - A^*\,] + 2r(A - A^2) - 2r(A) - 2r(I_m - A) \\ &= 2r[\, A, \ I_m - A^*\,] - 2m \leq 0. \end{aligned}$$

Because the rank of a matrix is nonnegative, the above inequality in fact implies that $r[\, A, \ I_m - A^*\,] = m$ and $AA^{\dagger}$ commutes with $(I_m - A)^{\dagger}(I_m - A)$ for any square matrix A. Based on this result, one can easily see that $AA^{\dagger}$ also commutes with $(\lambda I_m - A)^{\dagger}(\lambda I_m - A)$ for any $\lambda \neq 0$. However, it is curious that $AA^{\dagger}$ does not commute with $(I_m - A)(I_m - A)^{\dagger}$ in general. In fact, we find by (7.4) that

$$\begin{aligned} & r[\, AA^{\dagger}(I_m - A)(I_m - A)^{\dagger} - (I_m - A)(I_m - A)^{\dagger}AA^{\dagger}\,] \\ &= 2r[\, A, \ I_m - A\,] + 2r(A - A^*A) - 2r(A) - 2r(I_m - A) \\ &= 2m + 2r(A - A^*A) - 2r(A) - 2r(I_m - A) \\ &= 2r(A - A^*A) - 2r(A - A^2). \end{aligned}$$

Thus $AA^{\dagger}$ commutes with $(I_m - A)(I_m - A)^{\dagger}$ if and only if $r(A - A^*A) = r(A - A^2)$.

Next replacing A and B by $I_m + A$ and $I_m - A$ in (7.2), respectively, we can get

$$\begin{aligned} & r[\,(I_m + A)(I_m + A)^{\dagger}(I_m - A)^{\dagger}(I_m - A) - (I_m - A)^{\dagger}(I_m - A)(I_m + A)(I_m + A)^{\dagger}\,] \\ &= 2r[\, I_m + A, \ I_m - A^*\,] + 2r(I_m - A^2) - 2r(I_m + A) - 2r(I_m - A) \\ &= 2r[\, I_m + A, \ I_m - A^*\,] - 2m \leq 0. \end{aligned}$$

This inequality implies that $r[\, I_m + A, \ I_m - A^*\,] = m$ and $(I_m + A)(I_m + A)^{\dagger}$ commutes with $(I_m - A)^{\dagger}(I_m - A)$ for any square matrix A. By (7.4) we also find that

$$\begin{aligned} & r[\,(I_m + A)(I_m + A)^{\dagger}(I_m - A)(I_m - A)^{\dagger} - (I_m - A)(I_m - A)^{\dagger}(I_m + A)(I_m + A)^{\dagger}\,] \\ &= 2r[\, I_m + A, \ I_m - A\,] + 2r[(I_m - A^*)(I_m + A) - 2r(I_m + A) - 2r(I_m - A) \\ &= 2m + 2r[(I_m - A^*)(I_m + A)] - 2r(I_m + A) - 2r(I_m - A) \\ &= 2r[(I_m - A^*)(I_m + A)] - 2r(I_m - A^2). \end{aligned}$$

Thus $(I_m+A)(I_m+A)^\dagger$ commutes with $(I_m-A)(I_m-A)^\dagger$ if and only if $r[(I_m-A^*)(I_m+A)] = r(I_m-A^2)$.

A general result is that for any two polynomials $p(\lambda)$ and $q(\lambda)$ without common roots, we have according to (7.2) and (1.17)

$$
\begin{aligned}
& r[\, p(A)p^\dagger(A)q^\dagger(A)q(A) - q^\dagger(A)q(A)p(A)p^\dagger(A)\,] \\
& = 2r[\, p(A), \ q(A^*)\,] + 2r[p(A)q(A)] - 2r[p(A)] - 2r[q(A)] \\
& = 2r[\, p(A), \ q(A^*)\,] - 2m \le 0.
\end{aligned}
$$

This implies that $r[\, p(A), \ q(A^*)\,] = m$ and $[p(A)p^\dagger(A)][q^\dagger(A)q(A)] = [q^\dagger(A)q(A)][p(A)p^\dagger(A)]$, that is, $p(A)p^\dagger(A)$ commutes with $q^\dagger(A)q(A)$. On the other hand, the fact $r[\, p(A), \ q(A^*)\,] = m$ can also alternatively be stated that the for any square matrix A and any two polynomials $p(\lambda)$ and $q(\lambda)$ without common roots the Hermitian matrix $p(A)p(A^*) + q(A^*)q(A)$ is always positive definite.

Observe that the Moore-Penrose inverses of $[\, A, \ B\,]$ and $\begin{bmatrix} A \\ C \end{bmatrix}$ can be expressed as

$$
[\, A, \ B\,]^\dagger = [\, A, \ B\,]^* \, ([\, A, \ B\,][\, A, \ B\,]^*)^\dagger = \begin{bmatrix} A^*(\, AA^* + BB^*\,)^\dagger \\ B^*(\, AA^* + BB^*\,)^\dagger \end{bmatrix},
$$

$$
\begin{bmatrix} A \\ C \end{bmatrix}^\dagger = \left(\begin{bmatrix} A \\ C \end{bmatrix}^* \begin{bmatrix} A \\ C \end{bmatrix} \right)^\dagger \begin{bmatrix} A \\ C \end{bmatrix}^* = [\, (\, A^*A + C^*C\,)^\dagger A^*, \ (\, A^*A + C^*C\,)^\dagger C^*\,].
$$

Based on the two expressions we can find a series of rank equalities related to $[\, A, \ B\,]$ and $\begin{bmatrix} A \\ C \end{bmatrix}$ and their consequences.

Theorem 7.4. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{k\times m}$ be given. Then*

(a) $r[\, AA^*(\, AA^* + BB^*\,)^\dagger A - A\,] = r(A) + r(B) - r[\, A, \ B\,]$.

(b) $r[\, AA^\dagger(\, AA^\dagger + BB^\dagger\,)^\dagger A - A\,] = r(A) + r(B) - r[\, A, \ B\,]$.

(c) $r[\, A(\, A^*A + C^*C\,)^\dagger A^*A - A\,] = r(A) + r(C) - r\begin{bmatrix} A \\ C \end{bmatrix}$.

(d) $r[\, A(\, A^\dagger A + C^\dagger C\,)^\dagger A^\dagger A - A\,] = r(A) + r(C) - r\begin{bmatrix} A \\ C \end{bmatrix}$.

(e) $r[\, A^*(\, AA^* + BB^*\,)^\dagger B\,] = r(A) + r(B) - r[\, A, \ B\,]$.

(f) $r[\, A^\dagger(\, AA^\dagger + BB^\dagger\,)^\dagger B\,] = r(A) + r(B) - r[\, A, \ B\,]$.

(g) $r[\, A(\, A^*A + C^*C\,)^\dagger C^*\,] = r(A) + r(C) - r\begin{bmatrix} A \\ C \end{bmatrix}$.

(h) $r[\, A(\, A^\dagger A + C^\dagger C\,)^\dagger C^\dagger\,] = r(A) + r(C) - r\begin{bmatrix} A \\ C \end{bmatrix}$.

(i) *The following five statements are equivalent*:

(1) $AA^*(\, AA^* + BB^*\,)^\dagger A = A$.

(2) $AA^\dagger(\, AA^\dagger + BB^\dagger\,)^\dagger A = A$.

(3) $A^*(\, AA^* + BB^*\,)^\dagger B = 0$.

(4) $A^\dagger(\, AA^\dagger + BB^\dagger\,)^\dagger B = 0$.

(5) $r[\, A, \ B\,] = r(A) + r(B)$, *i.e.*, $R(A) \cap R(B) = \{0\}$.

(j) *The following five statements are equivalent*:

(1) $A(\, A^*A + C^*C\,)^\dagger A^*A = A$.

(2) $A(\, A^\dagger A + C^\dagger C\,)^\dagger A^\dagger A = A$.

(3) $A(\, A^*A + C^*C\,)^\dagger C^* = 0$.

(4) $A(\, A^\dagger A + C^\dagger C\,)^\dagger C^\dagger = 0$.

(5) $r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r(C)$ *i.e.*, $R(A^*) \cap R(C^*) = \{0\}$.

Proof. We only show Parts (a) and (b). Note that $R(A) \subseteq R(\, AA^* + BB^*\,)$. Thus we find by (1.7) that

$$
r[\, AA^*(\, AA^* + BB^*\,)^\dagger A - A\,] \ = \ r\begin{bmatrix} AA^* + BB^* & A \\ AA^* & A \end{bmatrix} - r(\, AA^* + BB^*\,)
$$

$$= r\begin{bmatrix} BB^* & 0 \\ 0 & A \end{bmatrix} - r[\,A,\ B\,] = r(A) + r(B) - r[\,A,\ B\,],$$

as required for Part (a). Similarly

$$\begin{aligned} r[\,AA^\dagger(\,AA^\dagger + BB^\dagger\,)^\dagger A - A\,] &= r\begin{bmatrix} AA^\dagger + BB^\dagger & A \\ AA^\dagger & A \end{bmatrix} - r(\,AA^\dagger + BB^\dagger\,) \\ &= r\begin{bmatrix} BB^\dagger & 0 \\ 0 & A \end{bmatrix} - r[\,A,\ B\,] = r(A) + r(B) - r[\,A,\ B\,], \end{aligned}$$

as required for Part (b). □

Theorem 7.5. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given. Then*

(a) *The rank of $BB^\dagger A - AC^\dagger C$ satisfies*

$$r(\,BB^\dagger A - AC^\dagger C\,) = r\begin{bmatrix} B^*A \\ C \end{bmatrix} + r[\,AC^*,\ B\,] - r(B) - r(C).$$

(b) $BB^\dagger A = AC^\dagger C \Leftrightarrow r\begin{bmatrix} B^*A \\ C \end{bmatrix} = r(C)$ *and* $r[\,AC^*,\ B\,] = r(B) \Leftrightarrow R(AC^*) \subseteq R(B)$ *and* $R(A^*B) \subseteq R(C^*)$.

(c) $BB^\dagger A - AC^\dagger C$ *is nonsingular* $\Leftrightarrow r\begin{bmatrix} B^*A \\ C \end{bmatrix} = r[\,AC^*,\ B\,] = r(B) + r(C) = m.$

Proof. Follows from (4.1) by noticing that both $BB^\dagger$ and $C^\dagger C$ are idempotent. □

It is well known that the matrix equation $BXC = A$ is solvable if and only if $BB^\dagger AC^\dagger C = A$. This leads us to consider the rank of $A - BB^\dagger AC^\dagger C$.

Theorem 7.6. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given. Then*

$$r(\,A - BB^\dagger AC^\dagger C\,) = r\begin{bmatrix} A & AC^* & B \\ B^*A & 0 & 0 \\ C & 0 & 0 \end{bmatrix} - r(B) - r(C), \tag{7.5}$$

and

$$r(\,2A - BB^\dagger A - AC^\dagger C\,) = r\begin{bmatrix} A & AC^* & B \\ B^*A & 0 & 0 \\ C & 0 & 0 \end{bmatrix} - r(B) - r(C). \tag{7.6}$$

In particular,

$$BB^\dagger AC^\dagger C = A \Leftrightarrow BB^\dagger A + AC^\dagger C = 2A \Leftrightarrow R(A) \subseteq R(B) \ \ \text{and} \ \ R(A^*) \subseteq R(C^*). \tag{7.7}$$

Proof. Applying (2.8) and the rank cancellation law (1.8) to $A - BB^\dagger AC^\dagger C$ produces

$$\begin{aligned} & r(\,A - BB^\dagger AC^\dagger C\,) \\ &= r\begin{bmatrix} B^*AC^* & B^*BB^* & 0 \\ C^*CC^* & 0 & C^*C \\ 0 & BB^* & -A \end{bmatrix} - r(B) - r(C) \\ &= r\begin{bmatrix} B^*AC^* & B^*B & 0 \\ CC^* & 0 & C \\ 0 & B & -A \end{bmatrix} - r(B) - r(C) = r\begin{bmatrix} 0 & 0 & B^*A \\ 0 & 0 & C \\ AC^* & B & -A \end{bmatrix} - r(B) - r(C), \end{aligned}$$

as required for (7.5). In the same way we can show (7.6). The result in (7.7) is well known. □

Theorem 7.7. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given, and let $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$. Then*

$$r(\,A - BB^\dagger A - AC^\dagger C\,) = r(M) + r(CA^*B) - r(B) - r(C), \tag{7.8}$$

that is, the block matrix M satisfies the rank equality

$$r(M) = r(B) + r(C) - r(CA^*B) + r(\,A - BB^\dagger A - AC^\dagger C\,). \tag{7.9}$$

Proof. Applying (2.2) and (1.8) to $A - BB^\dagger A - AC^\dagger C$ yields

$$\begin{aligned} r(\,A - BB^\dagger A - AC^\dagger C\,) &= r\begin{bmatrix} B^*BB^* & 0 & B^*A \\ 0 & C^*CC^* & C^*C \\ BB^* & AC^* & A \end{bmatrix} - r(B) - r(C) \\ &= r\begin{bmatrix} B^*B & 0 & B^*A \\ 0 & CC^* & C \\ B & AC^* & A \end{bmatrix} - r(B) - r(C) \\ &= r\begin{bmatrix} 0 & -B^*AC^* & 0 \\ 0 & 0 & C \\ B & 0 & A \end{bmatrix} - r(B) - r(C) \\ &= r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} + r(B^*AC^*) - r(B) - r(C), \end{aligned}$$

as required for (7.8). □

Theorem 7.8. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given, and let $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$. Then*

(a) *The rank of $A - A(E_BAF_C)^\dagger A$ satisfies*

$$r[\,A - A(E_BAF_C)^\dagger A\,] = r(A) + r(B) + r(C) - r(M), \tag{7.10}$$

that is,

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(A) + r(B) + r(C) - r[\,A - A(E_BAF_C)^\dagger A\,], \tag{7.11}$$

where $E_B = I - BB^\dagger$ and $F_C = I - C^\dagger C$.

(b) *In particular,*

$$(E_BAF_C)^\dagger \in \{A^-\} \tag{7.12}$$

holds if and only if

$$r(M) = r(A) + r(B) + r(C), \quad i.e., \quad R(A) \cap R(B) = \{0\} \text{ and } R(A^*) \cap R(C^*) = \{0\}. \tag{7.13}$$

(c) $r[\,A - A(E_BA)^\dagger A\,] = r(A) + r(B) - r[\,A,\ B\,]$.

(d) $r[\,A - A(AF_C)^\dagger A\,] = r(A) + r(C) - r\begin{bmatrix} A \\ C \end{bmatrix}$.

(e) $A(E_BA)^\dagger A = A \Leftrightarrow r[\,A,\ B\,] = r(A) + r(B)$, *i.e.*, $R(A) \cap R(B) = \{0\}$.

(f) $A(AF_C)^\dagger A = A \Leftrightarrow r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r(C)$, *i.e.*, $R(A^*) \cap R(C^*) = \{0\}$.

Proof. Let $N = E_BAF_C$. Then it is easy to verify that $N^*NN^* = N^*AN^*$. In that case, applying (2.1), and then (1.2) and (1.3) to $A - A(E_BAF_C)^\dagger A$ yields

$$\begin{aligned} r[\,A - A(E_BAF_C)^\dagger A\,] = r[\,A - AN^\dagger A\,] &= r\begin{bmatrix} N^*NN^* & N^*A \\ AN^* & A \end{bmatrix} - r(M) \\ &= r\begin{bmatrix} N^*NN^* - N^*AN^* & 0 \\ 0 & A \end{bmatrix} - r(N) \\ &= r\begin{bmatrix} 0 & 0 \\ 0 & A \end{bmatrix} - r(N) \\ &= r(A) - r(N) = r(A) + r(B) + r(C) - r(M), \end{aligned}$$

as required for (7.10). The equivalence of (7.12) and (7.13) follows immediately from (7.11). □

It is known that for any B and C, the matrix $(E_BAF_C)^\dagger$ is always an outer inverse of A (Greville [51]). Thus the rank formula (7.11) can also be derived from (5.6).

In the remainder of this chapter, we establish various rank equalities related to ranks of Moore-Penrose inverses of block matrices, and then present their consequences.

Theorem 7.9. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given. Then*

(a) $r\left([\, A,\ B\,]^\dagger - \begin{bmatrix} A^\dagger \\ B^\dagger \end{bmatrix} \right) = r[\, AA^*B,\ BB^*A\,]$.

(b) $r\left(\begin{bmatrix} A \\ C \end{bmatrix}^\dagger - [\, A^\dagger,\ C^\dagger\,] \right) = r\begin{bmatrix} AC^*C \\ CA^*A \end{bmatrix}$.

(c) $r\left([\, A,\ B\,]^\dagger[\, A,\ B\,] - \begin{bmatrix} A^\dagger \\ B^\dagger \end{bmatrix}[\, A,\ B\,] \right) = r[\, AA^*B,\ BB^*A\,]$.

(d) $r\left(\begin{bmatrix} A \\ C \end{bmatrix}\begin{bmatrix} A \\ C \end{bmatrix}^\dagger - \begin{bmatrix} A \\ C \end{bmatrix}[\, A^\dagger,\ C^\dagger\,] \right) = r\begin{bmatrix} AC^*C \\ CA^*A \end{bmatrix}$.

In particular,

(e) $[\, A,\ B\,]^\dagger = \begin{bmatrix} A^\dagger \\ B^\dagger \end{bmatrix} \Leftrightarrow [\, A,\ B\,]^\dagger[\, A,\ B\,] = \begin{bmatrix} A^\dagger \\ B^\dagger \end{bmatrix}[\, A,\ B\,] \Leftrightarrow A^*B = 0$.

(f) $\begin{bmatrix} A \\ C \end{bmatrix}^\dagger = [\, A^\dagger,\ C^\dagger\,] \Leftrightarrow \begin{bmatrix} A \\ C \end{bmatrix}\begin{bmatrix} A \\ C \end{bmatrix}^\dagger = \begin{bmatrix} A \\ C \end{bmatrix}[\, A^\dagger,\ C^\dagger\,] \Leftrightarrow CA^* = 0$.

Proof. Let $M = [\, A,\ B\,]$. Then it follows by (2.7) that

$$
\begin{aligned}
& r\left([\, A,\ B\,]^\dagger - \begin{bmatrix} A^\dagger \\ B^\dagger \end{bmatrix} \right) \\
= \ & r\left([\, A,\ B\,]^\dagger - \begin{bmatrix} I \\ 0 \end{bmatrix} A^\dagger - \begin{bmatrix} 0 \\ I \end{bmatrix} B^\dagger \right) \\
= \ & r\begin{bmatrix} -M^*MM^* & 0 & 0 & M^* \\ 0 & A^*AA^* & 0 & A^* \\ 0 & 0 & B^*BB^* & B^* \\ \begin{bmatrix} A^* \\ B^* \end{bmatrix} & \begin{bmatrix} A^* \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ B^* \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{bmatrix} - r(M) - r(A) - r(B) \\
= \ & r\begin{bmatrix} -M^*MM^* & 0 & 0 & M^* \\ -A^*AA^* & 0 & 0 & A^* \\ -B^*BB^* & 0 & 0 & B^* \\ 0 & A^* & 0 & 0 \\ 0 & 0 & B^* & 0 \end{bmatrix} - r(M) - r(A) - r(B) \\
= \ & r\begin{bmatrix} M^*MM^* & M^* \\ A^*AA^* & A^* \\ B^*BB^* & B^* \end{bmatrix} - r(M) \\
= \ & r\begin{bmatrix} MM^*M & AA^*A & BB^*B \\ M & A & B \end{bmatrix} - r(M) \\
= \ & r\begin{bmatrix} 0 & AA^*A - MM^*A & BB^*B - MM^*B \\ M & 0 & 0 \end{bmatrix} - r(M) \\
= \ & r[\, AA^*A - MM^*A,\ BB^*B - MM^*B\,] = r[\, AA^*B,\ BB^*A\,],
\end{aligned}
$$

as required in Part (a). Similarly, we can show Parts (b), (c) and (d). The results in Parts (e) and (f) follow immediately from Parts (a)—(d). $\square$

A general result is given below, the proof is omitted.

Theorem 7.10. *Let $A = [\,A_1,\ A_2,\ \cdots,\ A_k\,] \in \mathcal{C}^{m\times n}$ be given, and denote $M = \begin{bmatrix} A_1^\dagger \\ \vdots \\ A_k^\dagger \end{bmatrix}$. Then*

$$r(\,A^\dagger - M\,) = r(\,A^\dagger A - MA\,) = r[\,N_1N_1^*A_1,\ N_2N_2^*A_2,\ \cdots,\ N_kN_k^*A_k\,], \tag{7.14}$$

where $N_i = [\,A_1,\ \cdots,\ A_{i-1},\ A_{i+1},\ \cdots,\ A_k\,]$, $\ i = 1, 2, \cdots, k$. In particular,

$$A^\dagger = M \Leftrightarrow A^\dagger A = MA \Leftrightarrow A_iA_j^* = 0 \quad for\ all\ i \neq j. \tag{7.15}$$

Let $[\,A,\ B\,]^\dagger = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}$ and $\begin{bmatrix} A \\ C \end{bmatrix}^\dagger = [\,H_1,\ H_2\,]$. We next consider the relationships between A and G_1, A and H_1.

Theorem 7.11. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given. Then*

(a) $r(\,A^\dagger - [\,I_n,\ 0\,][\,A,\ B\,]^\dagger\,) = r(\,B^\dagger - [\,0,\ I_k\,][\,A,\ B\,]^\dagger\,) = r(A^*B)$.

(b) $r\left(A^\dagger - \begin{bmatrix} A \\ C \end{bmatrix}^\dagger \begin{bmatrix} I_m \\ 0 \end{bmatrix}\right) = r\left(C^\dagger - \begin{bmatrix} A \\ C \end{bmatrix}^\dagger \begin{bmatrix} 0 \\ I_l \end{bmatrix}\right) = r(CA^*)$.

(c) $r(\,AA^\dagger - [\,A,\ 0\,][\,A,\ B\,]^\dagger\,) = r(\,BB^\dagger - [\,0,\ B\,][\,A,\ B\,]^\dagger\,) = r(A^*B)$.

(d) $r\left(A^\dagger A - \begin{bmatrix} A \\ C \end{bmatrix}^\dagger \begin{bmatrix} A \\ 0 \end{bmatrix}\right) = r\left(C^\dagger C - \begin{bmatrix} A \\ C \end{bmatrix}^\dagger \begin{bmatrix} 0 \\ C \end{bmatrix}\right) = r(CA^*)$.

In particular,

(e) $[\,I_m,\ 0\,][\,A,\ B\,]^\dagger = A^\dagger \Leftrightarrow [\,0,\ I_k\,][\,A,\ B\,]^\dagger = B^\dagger \Leftrightarrow [\,A,\ 0\,][\,A,\ B\,]^\dagger = AA^\dagger \Leftrightarrow [\,0,\ B\,][\,A,\ B\,]^\dagger = B^\dagger B \Leftrightarrow A^*B = 0$.

(f) $\begin{bmatrix} A \\ C \end{bmatrix}^\dagger \begin{bmatrix} I_m \\ 0 \end{bmatrix} = A^\dagger \Leftrightarrow \begin{bmatrix} A \\ C \end{bmatrix}^\dagger \begin{bmatrix} 0 \\ I_l \end{bmatrix} = C^\dagger \Leftrightarrow \begin{bmatrix} A \\ C \end{bmatrix}^\dagger \begin{bmatrix} A \\ 0 \end{bmatrix} = A^\dagger A \Leftrightarrow \begin{bmatrix} A \\ C \end{bmatrix}^\dagger \begin{bmatrix} 0 \\ C \end{bmatrix} = C^\dagger C \Leftrightarrow CA^* = 0$.

Proof. We only prove Parts (a) and (c). Let $M = [\,A,\ B\,]$. Then it follows by (2.7), (1.8) and block elementary operations of matrices that

$$\begin{aligned} r(\,A^\dagger - [\,I_m,\ 0\,]M^\dagger\,) &= r\begin{bmatrix} -A^*AA^* & 0 & A^* \\ 0 & M^*MM^* & M^* \\ A^* & [\,I_m,\ 0\,]\,M^* & 0 \end{bmatrix} - r(A) - r(M) \\ &= r\begin{bmatrix} A^*AA^* & -A^*MM^* & A^* \\ 0 & 0 & M^* \\ A^* & A^* & 0 \end{bmatrix} - r(A) - r(M) \\ &= r\begin{bmatrix} 0 & -A^*BB^* & A^* \\ 0 & 0 & M^* \\ A^* & 0 & 0 \end{bmatrix} - r(A) - r(M) = r(A^*BB^*) = r(A^*B), \end{aligned}$$

establishing the first equality in Part (a). Similarly

$$\begin{aligned} r(\,AA^\dagger - [\,A,\ 0\,]M^\dagger\,) &= r\begin{bmatrix} -A^*AA^* & 0 & A^* \\ 0 & M^*MM^* & M^* \\ AA^* & [\,A,\ 0\,]\,M^* & 0 \end{bmatrix} - r(A) - r(M) \\ &= r\begin{bmatrix} A^*A & -A^*MM^* & A^* \\ 0 & 0 & M^* \\ A & AA^* & 0 \end{bmatrix} - r(A) - r(M) \\ &= r\begin{bmatrix} 0 & -A^*AA^* + A^*MM^* & A^* \\ 0 & 0 & M^* \\ A^* & 0 & 0 \end{bmatrix} - r(A) - r(M) \\ &= r\begin{bmatrix} 0 & A^*BB^* & A^* \\ 0 & 0 & A^* \\ 0 & 0 & B^* \\ A^* & 0 & 0 \end{bmatrix} - r(A) - r(M) = r(A^*BB^*) = r(A^*B), \end{aligned}$$

establishing the first equality in Part (c). $\square$

Theorem 7.12. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given. Then*

(a) $r(\, [A,\ B][A,\ B]^{\dagger} - (\, AA^{\dagger} + BB^{\dagger}\,)\,) = r[\,A,\ B\,] + 2r(A^*B) - r(A) - r(B)$.

(b) $r\left(\begin{bmatrix} A \\ C \end{bmatrix}^{\dagger} \begin{bmatrix} A \\ C \end{bmatrix} - (\, A^{\dagger}A + C^{\dagger}C\,) \right) = r\begin{bmatrix} A \\ C \end{bmatrix} + 2r(CA^*) - r(A) - r(C)$.

In particular,

(c) $[\,A,\ B\,][\,A,\ B\,]^{\dagger} = AA^{\dagger} + BB^{\dagger} \Leftrightarrow A^*B = 0 \Leftrightarrow [\,A,\ B\,]^{\dagger} = \begin{bmatrix} A^{\dagger} \\ B^{\dagger} \end{bmatrix}$.

(d) $\begin{bmatrix} A \\ C \end{bmatrix}^{\dagger} \begin{bmatrix} A \\ C \end{bmatrix} = A^{\dagger}A + C^{\dagger}C \Leftrightarrow CA^* = 0 \Leftrightarrow \begin{bmatrix} A \\ C \end{bmatrix}^{\dagger} = [\,A^{\dagger},\ C^{\dagger}\,]$.

Proof. Let $M = [\,A,\ B\,]$. Then it follows by (2.7), (1.8) and block elementary operations of matrices that

$$
\begin{aligned}
& r(\, MM^{\dagger} - AA^{\dagger} - BB^{\dagger}\,) \\
= \ & r\begin{bmatrix} -M^*MM^* & 0 & 0 & M^* \\ 0 & A^*AA^* & 0 & A^* \\ 0 & 0 & B^*BB^* & B^* \\ MM^* & AA^* & BB^* & 0 \end{bmatrix} - r(M) - r(A) - r(B) \\
= \ & r\begin{bmatrix} -M^*M & 0 & 0 & M^* \\ 0 & A^*A & 0 & A^* \\ 0 & 0 & B^*B & B^* \\ M & A & B & 0 \end{bmatrix} - r(M) - r(A) - r(B) \\
= \ & r\begin{bmatrix} 0 & 0 & 0 & 0 & A^* \\ 0 & 0 & 0 & 0 & B^* \\ A^*A & A^*B & A^*A & 0 & A^* \\ B^*A & B^*B & 0 & B^*B & B^* \\ A & B & A & B & 0 \end{bmatrix} - r(M) - r(A) - r(B) \\
= \ & r\begin{bmatrix} 0 & 0 & 0 & 0 & A^* \\ 0 & 0 & 0 & 0 & B^* \\ 0 & 0 & 0 & -A^*B & 0 \\ 0 & 0 & -B^*A & 0 & 0 \\ A & B & 0 & 0 & 0 \end{bmatrix} - r(M) - r(A) - r(B) \\
= \ & r(M) + 2r(A^*B) - r(A) - r(B),
\end{aligned}
$$

as required in Part (a). In the same way, we can show Part (b). We know from Part (a) that

$$
MM^{\dagger} = AA^{\dagger} + BB^{\dagger} \Leftrightarrow r[\,A,\ B\,] = r(A) + r(B) - 2r(A^*B). \tag{7.16}
$$

On the other hand, observe from (1.2) that

$$
\begin{aligned}
r[\,A,\ B\,] & = r(A) + r(\,B - AA^{\dagger}B\,) \\
& \geq r(A) + r(B) - r(AA^{\dagger}B) \\
& = r(A) + r(B) - r(A^*B) \\
& \geq r(A) + r(B) - 2r(A^*B).
\end{aligned}
$$

Thus (7.16) is also equivalent to $A^*B = 0$. In the similar manner, we can show Part (d). $\square$

A general result is given below, the proof is omitted.

Corollary 7.13. *Let $A = [\,A_1,\, A_2,\, \cdots,\, A_k\,] \in \mathcal{C}^{m\times n}$ be given. Then*

$$
r[AA^{\dagger} - (\, A_1A_1^{\dagger} + \cdots + A_kA_k^{\dagger}\,)] = r\begin{bmatrix} 0 & A_1^*A_2 & \cdots & A_1^*A_k \\ A_2^*A_1 & 0 & \cdots & A_2^*A_k \\ \vdots & \vdots & \ddots & \vdots \\ A_k^*A_1 & A_k^*A_2 & \cdots & 0 \end{bmatrix} + r(A) - r(A_1) - \cdots - r(A_k). \tag{7.17}
$$

In particular,

$$AA^{\dagger} = A_1A_1^{\dagger} + \cdots + A_kA_k^{\dagger} \Leftrightarrow A_i^*A_j = 0, \quad for\ all\ i \neq j. \tag{7.18}$$

Theorem 7.14. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given. Then*

(a) $r\left([\,A,\ B\,]^{\dagger} - \begin{bmatrix} (E_BA)^{\dagger} \\ (E_AB)^{\dagger} \end{bmatrix} \right) = r(A) + r(B) - r[\,A,\ B\,].$

(b) $r\left(\begin{bmatrix} A \\ C \end{bmatrix}^{\dagger} - [\,(AF_C)^{\dagger},\ (CF_A)^{\dagger}\,] \right) = r(A) + r(C) - r\begin{bmatrix} A \\ C \end{bmatrix}.$

In particular,

(c) $[\,A,\ B\,]^{\dagger} = \begin{bmatrix} (E_BA)^{\dagger} \\ (E_AB)^{\dagger} \end{bmatrix} \Leftrightarrow r[\,A,\ B\,] = r(A) + r(B),$ *i.e.,* $R(A) \cap R(B) = \{0\}.$

(d) $\begin{bmatrix} A \\ C \end{bmatrix}^{\dagger} = [\,(AF_C)^{\dagger},\ (CF_A)^{\dagger}\,] \Leftrightarrow r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r(C),$ *i.e.,* $R(A^*) \cap R(C^*) = \{0\}.$

Proof. Let $M = [\,A,\ B\,]$. Then it follows by (2.7) and (1.8) and block elementary operations of matrices that

$$\begin{aligned}
& r\left([\,A,\ B\,]^{\dagger} - \begin{bmatrix} (E_BA)^{\dagger} \\ (E_AB)^{\dagger} \end{bmatrix} \right) \\
= \ & r\left([\,A,\ B\,]^{\dagger} - \begin{bmatrix} I \\ 0 \end{bmatrix}(E_BA)^{\dagger} - \begin{bmatrix} 0 \\ I \end{bmatrix}(E_AB)^{\dagger} \right) \\
= \ & r\begin{bmatrix} -M^*MM^* & 0 & 0 & M^* \\ 0 & (E_BA)^*(E_BA)(E_BA)^* & 0 & (E_BA)^* \\ 0 & 0 & (E_AB)^*(E_AB)(E_AB)^* & (E_AB)^* \\ \begin{bmatrix} A^* \\ B^* \end{bmatrix} & \begin{bmatrix} (E_BA)^* \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ (E_AB)^* \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{bmatrix} \\
& -r(M) - r(E_BA) - r(E_AB) \\
= \ & r\begin{bmatrix} -M^*MM^* & 0 & 0 & M^* \\ 0 & (E_BA)^*A(E_BA)^* & 0 & (E_BA)^* \\ 0 & 0 & (E_AB)^*B(E_AB)^* & (E_AB)^* \\ A^* & (E_BA)^* & 0 & 0 \\ B^* & 0 & (E_AB)^* & 0 \end{bmatrix} \\
& -r(M) - r(E_BA) - r(E_AB) \\
= \ & r\begin{bmatrix} 0 & M^*A(E_BA)^* & M^*B(E_AB)^* & M^* \\ 0 & (E_BA)^*A(E_BA)^* & 0 & (E_BA)^* \\ 0 & 0 & (E_AB)^*B(E_AB)^* & (E_AB)^* \\ A^* & 0 & 0 & 0 \\ B^* & 0 & 0 & 0 \end{bmatrix} - r(M) - r(E_BA) - r(E_AB) \\
= \ & r\begin{bmatrix} M^*A(E_BA)^* & M^*B(E_AB)^* & M^* \\ (E_BA)^*A(E_BA)^* & 0 & (E_BA)^* \\ 0 & (E_AB)^*B(E_AB)^* & (E_AB)^* \end{bmatrix} - r(E_BA) - r(E_AB) \\
= \ & r\begin{bmatrix} 0 & 0 & M^* \\ 0 & 0 & (E_BA)^* \\ 0 & 0 & (E_AB)^* \end{bmatrix} - r(E_BA) - r(E_AB) \\
= \ & r[\,M,\ E_BA,\ E_AB\,] - r(E_BA) - r(E_AB) \\
= \ & r[\,A,\ B\,] - r(E_BA) - r(E_AB) = r(A) + r(B) - r[\,A,\ B\,],
\end{aligned}$$

as required for Part (a). Similarly, we can show Part (b). The results in Parts (c) and (d) follow immediately from Parts (a) and (b). $\square$

A general result is given below, its proof is much similar to that of Theorem 7.14 and is, therefore, omitted.

Theorem 7.15. *Let* $A = [\, A_1, \ A_2, \ \cdots, \ A_k \,] \in \mathcal{C}^{m \times n}$ *be given. Then*

$$r\left([\, A_1, \ A_2, \ \cdots, \ A_k \,]^\dagger - \begin{bmatrix} (E_{N_1}A_1)^\dagger \\ (E_{N_2}A_2)^\dagger \\ \vdots \\ (E_{N_k}A_k)^\dagger \end{bmatrix} \right) = r(A_1) + r(A_2) + \cdots + r(A_k) - r(A), \tag{7.19}$$

where $N_i = [\, A_1, \ \cdots, \ A_{i-1}, \ A_{i+1}, \cdots, \ A_k \,], \ i = 1, \ 2, \ \cdots, \ k$. *In particular,*

$$[\, A_1, \ A_2, \ \cdots, \ A_k \,]^\dagger = \begin{bmatrix} (E_{N_1}A_1)^\dagger \\ (E_{N_2}A_2)^\dagger \\ \vdots \\ (E_{N_k}A_k)^\dagger \end{bmatrix} \Leftrightarrow r(A) = r(A_1) + r(A_2) + \cdots + r(A_k). \tag{7.20}$$

Theorem 7.16. *Let* $A \in \mathcal{C}^{m \times n}, \ B \in \mathcal{C}^{m \times k}$ *and* $C \in \mathcal{C}^{l \times n}$ *be given. Then*

(a) $r\left([\, A, \ B \,][\, A, \ B \,]^\dagger - A(E_BA)^\dagger - B(E_AB)^\dagger \right) = r(A) + r(B) - r[\, A, \ B \,]$.

(b) $r\left(\begin{bmatrix} A \\ C \end{bmatrix}^\dagger \begin{bmatrix} A \\ C \end{bmatrix} - (AF_C)^\dagger A - (CF_A)^\dagger C \right) = r(A) + r(C) - r\begin{bmatrix} A \\ C \end{bmatrix}$.

In particular,

(c) $[\, A, \ B \,][\, A, \ B \,]^\dagger = A(E_BA)^\dagger + A(E_AB)^\dagger \Leftrightarrow R(A) \cap R(B) = \{0\}$.

(d) $\begin{bmatrix} A \\ C \end{bmatrix}^\dagger \begin{bmatrix} A \\ C \end{bmatrix} = (AF_C)^\dagger A + (CF_A)^\dagger C \Leftrightarrow R(A^*) \cap R(C^*) = \{0\}$.

The proof of Theorem 7.16 is much similar to that of Theorem 7.14 and is, therefore, omitted.

Theorem 7.17. *Let* $A \in \mathcal{C}^{m \times n}$, $B \in \mathcal{C}^{m \times k}$ *and* $C \in \mathcal{C}^{l \times n}$ *be given. Then*

(a) $r\left([\, A, \ B \,]^\dagger [\, A, \ B \,] - \begin{bmatrix} A^\dagger A & 0 \\ 0 & B^\dagger B \end{bmatrix} \right) = r(A) + r(B) - r[\, A, \ B \,]$.

(b) $r\left([\, A, \ B \,]^\dagger [\, A, \ B \,] - \begin{bmatrix} (E_BA)^\dagger(E_BA) & 0 \\ 0 & (E_AB)^\dagger(E_AB) \end{bmatrix} \right) = r(A) + r(B) - r[\, A, \ B \,]$.

(c) $r\left(\begin{bmatrix} A \\ C \end{bmatrix} \begin{bmatrix} A \\ C \end{bmatrix}^\dagger - \begin{bmatrix} AA^\dagger & 0 \\ 0 & CC^\dagger \end{bmatrix} \right) = r(A) + r(C) - r\begin{bmatrix} A \\ C \end{bmatrix}$.

(d) $r\left(\begin{bmatrix} A \\ C \end{bmatrix} \begin{bmatrix} A \\ C \end{bmatrix}^\dagger - \begin{bmatrix} (AF_C)(AF_C)^\dagger & 0 \\ 0 & (CF_A)(CF_A)^\dagger \end{bmatrix} \right) = r(A) + r(C) - r\begin{bmatrix} A \\ C \end{bmatrix}$.

In particular,

(e) $[\, A, \ B \,]^\dagger [\, A, \ B \,] = \begin{bmatrix} A^\dagger A & 0 \\ 0 & B^\dagger B \end{bmatrix} \Leftrightarrow [\, A, \ B \,]^\dagger [\, A, \ B \,] = \begin{bmatrix} (E_BA)^\dagger(E_BA) & 0 \\ 0 & (E_AB)^\dagger(E_AB) \end{bmatrix} \Leftrightarrow$ $R(A) \cap R(B) = \{0\}$.

(f) $\begin{bmatrix} A \\ C \end{bmatrix} \begin{bmatrix} A \\ C \end{bmatrix}^\dagger = \begin{bmatrix} AA^\dagger & 0 \\ 0 & CC^\dagger \end{bmatrix} \Leftrightarrow \begin{bmatrix} A \\ C \end{bmatrix} \begin{bmatrix} A \\ C \end{bmatrix}^\dagger = \begin{bmatrix} (AF_C)(AF_C)^\dagger & 0 \\ 0 & (CF_A)(CF_A)^\dagger \end{bmatrix} \Leftrightarrow R(A^*) \cap$ $R(C^*) = \{0\}$.

Proof. Let $M = [\, A, \ B \,]$ and $N = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$. Then we find by Theorem 7.2(b) that

$$\begin{aligned} r\left([\, A, \ B \,]^\dagger [\, A, \ B \,] - \begin{bmatrix} A^\dagger A & 0 \\ 0 & B^\dagger B \end{bmatrix} \right) &= r(\, M^\dagger M - N^\dagger N \,) \\ &= 2r\begin{bmatrix} M \\ N \end{bmatrix} - r(M) - r(N) \\ &= 2r\begin{bmatrix} A & B \\ A & 0 \\ 0 & B \end{bmatrix} - r[\, A, \ B \,] - r(A) - r(B) \\ &= r(A) + r(B) - r[\, A, \ B \,], \end{aligned}$$

as required for Part (a). Similarly we can show Parts (b)—(d). Parts (e) and (f) are direct consequences of Parts (a) and (b). $\square$

A general result is given below, and its proof is omitted.

Corollary 7.18. *Let $A = [\, A_1, \, A_2, \, \cdots, \, A_k \,] \in \mathcal{C}^{m \times n}$ be given. Then*

$$r[\, A^{\dagger}A - \mathrm{diag}(\, A_1^{\dagger}A_1, \;\, A_2^{\dagger}A_2, \;\, \cdots, \;\, A_k^{\dagger}A_k \,)\,] = r(A_1) + r(A_2) + \cdots + r(A_k) - r(A). \tag{7.21}$$

In particular,

$$A^{\dagger}A = \mathrm{diag}(\, A_1^{\dagger}A_1, \;\, A_2^{\dagger}A_2, \;\, \cdots, \;\, A_k^{\dagger}A_k \,) \Leftrightarrow r(A) = r(A_1) + r(A_2) + \cdots + r(A_k). \tag{7.22}$$

Theorem 7.19. *Let $A \in \mathcal{C}^{m \times n}$, $B \in \mathcal{C}^{m \times k}$ and $C \in \mathcal{C}^{l \times n}$ be given, and let $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$. Then*

$$r\left(A - [\, A, \; 0\,]M^{\dagger} \begin{bmatrix} A \\ 0 \end{bmatrix} \right) = r(A) + r(B) + r(C) - r(M), \tag{7.23}$$

or alternatively

$$r(M) = r(A) + r(B) + r(C) - r\left(A - [\, A, \; 0\,]M^{\dagger} \begin{bmatrix} A \\ 0 \end{bmatrix} \right). \tag{7.24}$$

In particular,

$$A[\, I_n, \; 0\,] \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^{\dagger} \begin{bmatrix} I_m \\ 0 \end{bmatrix} A = A \tag{7.25}$$

holds if and only if

$$r(M) = r(A) + r(B) + r(C), \quad i.e., \quad R(A) \cap R(B) = \{0\} \;\; and \;\; R(A^*) \cap R(C^*) = \{0\}. \tag{7.26}$$

Proof. It follows by (2.1) that

$$\begin{aligned} r\left(A - [\, A, \; 0\,]M^{\dagger} \begin{bmatrix} A \\ 0 \end{bmatrix} \right) &= r \begin{bmatrix} M^*MM^* & M^* \begin{bmatrix} A \\ 0 \end{bmatrix} \\ [\, A, \; 0\,]\, M^* & A \end{bmatrix} - r(M) \\ &= r \begin{bmatrix} M^*MM^* - M^* \begin{bmatrix} I \\ 0 \end{bmatrix} A\,[\, I, \; 0\,]\, M^* & 0 \\ 0 & A \end{bmatrix} - r(M) \\ &= r \left(M^* \begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix} M^* \right) + r(A) - r(M) \\ &= r \begin{bmatrix} BB^*A + AC^*C & BB^*B \\ CC^*C & 0 \end{bmatrix} + r(A) - r(M) \\ &= r(A) + r(B) + r(C) - r(M), \end{aligned}$$

as required for (7.23). $\square$

Theorem 7.20. *Let $A \in \mathcal{C}^{m \times n}$, $B \in \mathcal{C}^{m \times k}$ and $C \in \mathcal{C}^{l \times n}$ be given, and let $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$. Then*

$$r\left((E_BAF_C)^{\dagger} - [\, I_n, \; 0\,]M^{\dagger} \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right) = r \begin{bmatrix} A \\ C \end{bmatrix} + r[\, A, \;\, B\,] + r(B) + r(C) - 2r(M), \tag{7.27}$$

or alternatively

$$r(M) = \frac{1}{2}r \begin{bmatrix} A \\ C \end{bmatrix} + \frac{1}{2}r[\, A, \;\, B\,] + \frac{1}{2}r(B) + \frac{1}{2}r(C) - \frac{1}{2}r\left((E_BAF_C)^{\dagger} - [\, I_n, \; 0\,]M^{\dagger} \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right). \tag{7.28}$$

In particular,

$$[\, I_n, \; 0\,]M^{\dagger} \begin{bmatrix} I_m \\ 0 \end{bmatrix} = (E_BAF_C)^{\dagger} \Leftrightarrow r(M) = r \begin{bmatrix} A \\ C \end{bmatrix} + r(B) = r[\, A, \;\, B\,] + r(C). \tag{7.29}$$

Proof. Let $P = [\, I_n, \ 0 \,]$ and $Q = \begin{bmatrix} I_m \\ 0 \end{bmatrix}$. It follows by (2.2) that

$$\begin{aligned}
& r\left[\, (E_BAF_C)^\dagger - PM^\dagger Q \right] \\
&= r\begin{bmatrix} (E_BAF_C)^*A(E_BAF_C)^* & 0 & (E_BAF_C)^* \\ 0 & -M^*MM^* & M^*Q \\ (E_BAF_C)^* & PM^* & 0 \end{bmatrix} - r(E_BAF_C) - r(M) \\
&= r\begin{bmatrix} 0 & 0 & (E_BAF_C)^* \\ 0 & -M^*MM^* + M^*QAPM^* & M^*Q \\ (E_BAF_C)^* & PM^* & 0 \end{bmatrix} + r(B) + r(C) - 2r(M) \\
&= r\begin{bmatrix} 0 & 0 & (E_BAF_C)^* \\ 0 & M^*\begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix}M^* & M^*Q \\ (E_BAF_C)^* & PM^* & 0 \end{bmatrix} + r(B) + r(C) - 2r(M) \\
&= r\begin{bmatrix} 0 & 0 & 0 & E_BAF_C \\ 0 & 0 & 0 & A \\ 0 & 0 & 0 & C \\ E_BAF_C & A & B & 0 \end{bmatrix} + r(B) + r(C) - 2r(M) \\
&= r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A, \ B\,] + r(B) + r(C) - 2r(M),
\end{aligned}$$

establishing (7.27). $\square$

Theorem 7.21. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given. Then*

$$r\left(A + [\,0, \ B\,]\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^\dagger \begin{bmatrix} 0 \\ C \end{bmatrix}\right) = r[\,A, \ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}, \tag{7.30}$$

or alternatively

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r[\,A, \ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\left(A + [\,0, \ B\,]\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^\dagger \begin{bmatrix} 0 \\ C \end{bmatrix}\right). \tag{7.31}$$

Proof. Let $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$. Then it follows by (2.1) and block elementary operation that

$$\begin{aligned}
r\left(A + [\,0, \ B\,]M^\dagger \begin{bmatrix} 0 \\ C \end{bmatrix}\right) &= r\begin{bmatrix} M^*MM^* & M^*\begin{bmatrix} 0 \\ C \end{bmatrix} \\ [\,0, \ B\,]\,M^* & -A \end{bmatrix} - r(M) \\
&= r\begin{bmatrix} M^*MM^* - M^*\begin{bmatrix} 0 \\ I \end{bmatrix}C\,[\,I, \ 0\,]\,M^* & M^*\begin{bmatrix} 0 \\ C \end{bmatrix} \\ [\,A, \ B\,]\,M^* & -A \end{bmatrix} - r(M) \\
&= r\begin{bmatrix} M^*MM^* - M^*\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}M^* & M^*\begin{bmatrix} A \\ C \end{bmatrix} \\ [\,A, \ B\,]\,M^* & -A \end{bmatrix} - r(M) \\
&= r\begin{bmatrix} 0 & M^*\begin{bmatrix} A \\ C \end{bmatrix} \\ [\,A, \ B\,]\,M^* & -A \end{bmatrix} - r(M) \\
&= r\begin{bmatrix} 0 & \begin{bmatrix} A \\ C \end{bmatrix} \\ [\,A, \ B\,] & -A \end{bmatrix} - r(M) = r[\,A, \ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r(M),
\end{aligned}$$

as requird for (7.30). $\square$

It is easy to derive from (1.6) that

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} \geq r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,] - r(A).$$

Now replacing A by $A - BXC$ in the above inequality, where X is arbitrary, we obtain

$$r(\,A - BXC\,) \geq r[\,A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}. \tag{7.32}$$

This rank inequality implies that the quantity in the right-hand side of (7.32) is a lower bound for the rank of $A - BXC$ with respect to the choice of X. Combining (7.30) and (7.32), we immediately obtain

$$\min_X r(\,A - BXC\,) = r[\,A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}, \tag{7.33}$$

and a matrix satisfying (7.33) is given by

$$X = -[\,0,\ I_l\,]\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^{\dagger}\begin{bmatrix} 0 \\ I_k \end{bmatrix}. \tag{7.34}$$

Theorem 7.22. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $D \in \mathcal{C}^{l\times k}$ be given. Then*

(a) $r\left(\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}^{\dagger} - \begin{bmatrix} AA^{\dagger} & 0 \\ 0 & DD^{\dagger} \end{bmatrix}\right) = r(D) - r(A) + 2r[\,A,\ B\,] - r\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}.$

(b) $r\left(\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}^{\dagger}\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} - \begin{bmatrix} A^{\dagger}A & 0 \\ 0 & D^{\dagger}D \end{bmatrix}\right) = r(A) - r(D) + 2r\begin{bmatrix} B \\ D \end{bmatrix} - r\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}.$

In particular,

(c) $\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}^{\dagger} = \begin{bmatrix} AA^{\dagger} & 0 \\ 0 & DD^{\dagger} \end{bmatrix} \Leftrightarrow R(B) \subseteq R(A).$

(d) $\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}^{\dagger}\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = \begin{bmatrix} A^{\dagger}A & 0 \\ 0 & D^{\dagger}D \end{bmatrix} \Leftrightarrow R(B^*) \subseteq R(D^*).$

Proof. Let $M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$ and $N = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$. Then we find by Theorem 7.2(a) that

$$\begin{aligned} r(\,MM^{\dagger} - NN^{\dagger}\,) &= 2r[\,M,\ N\,] - r(M) - r(N) \\ &= 2r\begin{bmatrix} A & B & A & 0 \\ 0 & D & 0 & D \end{bmatrix} - r(M) - r(A) - r(D) \\ &= 2r[\,A,\ B\,] + r(D) - r(M) - r(A), \end{aligned}$$

as required for Part (a). Similarly we can show Part (b). Observe that

$$r(D) - r(A) + 2r[\,A,\ B\,] - r\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = (r[\,A,\ B\,] - r(A)) + \left(r(D) + r[\,A,\ B\,] - r\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}\right),$$

and

$$r(A) - r(D) + 2r\begin{bmatrix} B \\ D \end{bmatrix} - r\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = \left(r\begin{bmatrix} B \\ D \end{bmatrix} - r(D)\right) + \left(r(A) + r\begin{bmatrix} B \\ D \end{bmatrix} - r\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}\right).$$

Thus Parts (c) and (d) follow. $\square$

A general result is given below, and the proof is omitted for simplicity.

Theorem 7.23. *Let*

$$M = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ & A_{22} & \cdots & A_{2k} \\ & & \ddots & \vdots \\ & & & A_{kk} \end{bmatrix} \in \mathcal{C}^{m\times n}$$

be given, and let $A = \mathrm{diag}(\, A_{11},\, A_{22},\, \cdots,\, A_{kk}\,)$. *Then*

(a) $r\left[\, MM^{\dagger} - \mathrm{diag}(\, A_{11}A_{11}^{\dagger},\ \cdots,\ A_{kk}A_{kk}^{\dagger}\,)\right] = 2r[\, M,\ A\,] - r(M) - r(A).$

(b) $r\left[\, M^{\dagger}M - \mathrm{diag}(\, A_{11}^{\dagger}A_{11},\ \cdots,\ A_{kk}^{\dagger}A_{kk}\,)\right] = 2r\begin{bmatrix} M \\ A \end{bmatrix} - r(M) - r(A).$

(c) $MM^{\dagger} = \mathrm{diag}(\, A_{11}A_{11}^{\dagger},\, \cdots,\ A_{kk}A_{kk}^{\dagger}\,) \Leftrightarrow R(M) = R(A) \Leftrightarrow R(A_{ij}) \subseteq R(A_{ii}),\ j = i+1,\, \cdots,\, k,\ i = 1,\, \cdots,\, k-1.$

(d) $M^{\dagger}M = \mathrm{diag}(\, A_{11}^{\dagger}A_{11},\ \cdots,\ A_{kk}^{\dagger}A_{kk}\,) \Leftrightarrow R(M^*) = R(A^*) \Leftrightarrow R(A_{ij}) \subseteq R(A_{jj}),\ j = 2,\ \cdots,\ k,\ i = 1,\ \cdots,\ j-1.$

By Theorem 7.2(a) and (b), we can also establish the following.

Theorem 7.24. *Let* $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ *and* $D \in \mathcal{C}^{l\times k}$ *be given. Then*

(a) $r\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger} - \begin{bmatrix} AA^{\dagger} & 0 \\ 0 & DD^{\dagger} \end{bmatrix}\right) = 2r[\, A,\ B\,] + 2r[\, C,\ D\,] - r(A) - r(D) - r\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$

(b) $r\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger}\begin{bmatrix} A & B \\ C & D \end{bmatrix} - \begin{bmatrix} A^{\dagger}A & 0 \\ 0 & D^{\dagger}D \end{bmatrix}\right) = 2r\begin{bmatrix} A \\ C \end{bmatrix} + 2r\begin{bmatrix} B \\ D \end{bmatrix} - r(A) - r(D) - r\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$

In particular,

(c) $\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger} = \begin{bmatrix} AA^{\dagger} & 0 \\ 0 & DD^{\dagger} \end{bmatrix} \Leftrightarrow r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r(A) + r(D),\ R(B) \subseteq R(A)$ *and* $R(C) \subseteq R(D)$.

(d) $\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger}\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A^{\dagger}A & 0 \\ 0 & D^{\dagger}D \end{bmatrix} \Leftrightarrow r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r(A) + r(D),\ R(C^*) \subseteq R(A^*)$ *and* $R(B^*) \subseteq R(D^*)$.

Chapter 8

Reverse order laws for Moore-Penrose inverses

Reverse order laws for generalized inverses of products of matrices have been an attractive topic in the theory of generalized inverses of matrices, for these laws can reveal essential relationships between generalized inverses of products of matrices and generalized inverses of each matrix in the products. Various results on reverse order laws related to inner inverses, reflexive inner inverses, Moore-Penrose inverses, group inverses, Drazin inverses, and weighted Moore-Penrose inverses of products of matrices have widely been established by lot of authors (see, e.g., [12, 13, 14, 30, 41, 43, 49, 50, 60, 123, 124, 130, 133, 135, 149, 150, 151]). In this chapter, we shall present some rank equalities related to products of Moore-Penrose inverses of matrices, and then derive from them various types of reverse order laws for Moore-Penrose inverses of products of matrices.

Theorem 8.1. *Let $A \in \mathcal{C}^{m\times n}$ and $B \in \mathcal{C}^{n\times p}$ be given. Then*

$$r(\, AB - ABB^{\dagger}A^{\dagger}AB \,) = r(\, B^{\dagger}A^{\dagger} - B^{\dagger}A^{\dagger}ABB^{\dagger}A^{\dagger} \,) = r[\, A^*, \ B \,] + r(AB) - r(A) - r(B). \tag{8.1}$$

In particular, the following seven statements are equivalent:

(a) $B^{\dagger}A^{\dagger} \in \{(AB)^-_r\}$, *i.e.*, $B^{\dagger}A^{\dagger}$ *is a reflexive inner inverse of* AB.
(b) $r[\, A^*, \ B \,] = r(A) + r(B) - r(AB)$.
(c) $\dim[R(A) \cap R(B^*)] = r(AB)$.
(d) $r(\, B - A^{\dagger}AB \,) = r(B) - r(A^{\dagger}AB)$, *i.e.*, $A^{\dagger}AB \leq_{rs} B$.
(e) $r(\, A - ABB^{\dagger} \,) = r(A) - r(ABB^{\dagger})$, *i.e.*, $ABB^{\dagger} \leq_{rs} A$.
(f)[64] $B^{\dagger}A^{\dagger}ABB^{\dagger}A^{\dagger} = B^{\dagger}A^{\dagger}$.
(g)[64] $AA^{\dagger}B^{\dagger}B = B^{\dagger}BAA^{\dagger}$.

Proof. Applying (2.8) and (1.7) to $AB - ABB^{\dagger}A^{\dagger}AB$, we obtain

$$\begin{aligned}
r(\, AB - ABB^{\dagger}A^{\dagger}AB \,) &= r\begin{bmatrix} B^*A^* & B^*BB^* & 0 \\ A^*AA^* & 0 & A^*AB \\ 0 & ABB^* & -AB \end{bmatrix} - r(A) - r(B) \\
&= r\begin{bmatrix} B^*A^* & B^*B & 0 \\ AA^* & 0 & AB \\ 0 & AB & -AB \end{bmatrix} - r(A) - r(B) \\
&= r\begin{bmatrix} B^*A^* & B^*B \\ AA^* & AB \end{bmatrix} + r(AB) - r(A) - r(B) \\
&= r\left(\begin{bmatrix} A \\ B^* \end{bmatrix} [\, A^*, \ B \,] \right) + r(AB) - r(A) - r(B) \\
&= r(\, [\, A^*, \ B \,]^*[\, A^*, \ B \,] \,) + r(AB) - r(A) - r(B) \\
&= r[\, A^*, \ B \,] + r(AB) - r(A) - r(B).
\end{aligned}$$

Thus we have the first part of (8.1). Replace A by $B^\dagger$ and B by $A^\dagger$ and simplify to yield the second part of (8.1). The equivalence of Parts (a), (b) and (f) follows immediately from (8.1). The equivalence of Parts (b) and (c) follows from the well-known rank formula

$$r[\, A^*, \ B\,] = r(A) + r(B) - \dim[R(A^*) \cap R(B)].$$

The equivalence of Parts (b), (d) and (e) follows from (1.2) and (1.3). The equivalence of Parts (b) and (g) follows from (7.2). $\square$

The rank formula (8.1) was established by Baksalary and Styan [9] in an alternative form

$$r(AE_BF_AB) = r[\, A^*, \ B\,] + r(AB) - r(A) - r(B). \tag{8.1$'$}$$

Observe that

$$A(\, I - BB^\dagger\,)(\, I - A^\dagger A\,)B = -AB + ABB^\dagger A^\dagger AB.$$

Thus (8.1$'$) is exactly (8.1). Some extensions and applications of (8.1$'$) in mathematical statistics were also considered by Baksalary and Styan [9]. But in this monograph we only consider the application of (8.1) to the reverse order law $B^\dagger A^\dagger \in \{(AB)^-\}$. In addition, the results in Theorem 8.1 can also be extended to a product of n matrices. The corresponding results were presented by the author in [135].

As an application of (8.1), we let $B = I_m - A$ in (8.1). Then

$$\begin{aligned}
& r[\, (A - A^2) - (A - A^2)(I - A)^\dagger A^\dagger (A - A^2)\,] \\
& = r[\, A^*, \ I_m - A\,] + r(A - A^2) - r(A) - r(I_m - A) \\
& = r[\, A^*, \ I_m - A\,] - m \le 0.
\end{aligned}$$

This inequality implies that $r[\, A^*, \ I_m - A\,] = m$ and $(I - A)^\dagger A^\dagger$ is a reflexive inner inverse of the matrix $A - A^2$. By symmetry, $(I - A)^\dagger A^\dagger$ is also a reflexive inner inverse of the matrix $A - A^2$.

Replacing A and B in (8.1) by $I_m + A$ and $I_m - A$, respectively, we then get

$$\begin{aligned}
& r[\, (I_m - A^2) - (I_m - A^2)(I - A)^\dagger (I_m + A)^\dagger (I_m - A^2)\,] \\
& = r[\, I_m + A^*, \ I_m - A\,] + r(I_m - A^2) - r(I_m + A) - r(I_m - A) \\
& = r[\, I_m + A^*, \ I_m - A\,] - m \le 0.
\end{aligned}$$

This inequality implies that $r[\, I_m + A^*, \ I_m - A\,] = m$ and $(I - A)^\dagger (I_m + A)^\dagger$ is a reflexive inner inverse of the matrix $I_m - A^2$. By symmetry, $(I + A)^\dagger (I_m - A)^\dagger$ is also a reflexive inner inverse of the matrix $I_m - A^2$.

In general, for any two polynomials $p(\lambda)$ and $q(\lambda)$ without common roots, we find by (8.1) and (1.17) the following

$$\begin{aligned}
& r[\, p(A)q(A) - p(A)q(A)q^\dagger(A)p^\dagger(A)p(A)q(A)\,] \\
& = r[\, p(A^*), \ q(A)\,] + r[p(A)q(A)] - r[p(A)] - r[q(A)] \\
& = r[\, p(A^*), \ q(A)\,] - m \le 0.
\end{aligned}$$

This implies that $r[\, p(A^*), \ q(A)\,] = m$ and $q^\dagger(A)p^\dagger(A)$ is a reflexive inner inverse of the matrix $p(A)q(A)$. By symmetry, $p^\dagger(A)q^\dagger(A)$ is also a reflexive inner inverse of the matrix $p(A)q(A)$.

Theorem 8.2. *Let $A \in \mathcal{C}^{m\times n}$ and $B \in \mathcal{C}^{n \times p}$ be given. Then*

(a) $r[\, (AB)(AB)^\dagger - (AB)(B^\dagger A^\dagger)\,] = r[\, B, \ A^*AB\,] - r(B) = r(\, A^*AB - BB^\dagger A^*AB\,)$.

(b) $r[\, (AB)^\dagger (AB) - (B^\dagger A^\dagger)(AB)\,] = r\begin{bmatrix} A \\ ABB^* \end{bmatrix} - r(A) = r(\, ABB^* - ABB^*A^\dagger A\,)$.

In particular,

(c) $(AB)(AB)^\dagger = (AB)(B^\dagger A^\dagger) \Leftrightarrow A^*AB = BB^\dagger A^*AB \Leftrightarrow R(A^*AB) \subseteq R(B) \Leftrightarrow B^\dagger A^\dagger \subseteq \{\, (AB)^{(1,2,3)}\}$.

(d) $(AB)^\dagger(AB) = (B^\dagger A^\dagger)(AB) \Leftrightarrow ABB^* = ABB^*A^\dagger A \Leftrightarrow R(BB^*A^*) \subseteq R(A^*) \Leftrightarrow B^\dagger A^\dagger \subseteq \{\, (AB)^{(1,2,4)}\}$.

(e) *The following four statements are equivalent*:

(1) $(AB)^\dagger = B^\dagger A^\dagger$.
(2) $(AB)(AB)^\dagger = (AB)(B^\dagger A^\dagger)$ *and* $(AB)^\dagger(AB) = (B^\dagger A^\dagger)(AB)$.
(3) $A^*AB = BB^\dagger A^*AB$ *and* $ABB^* = ABB^*A^\dagger A$.
(4) $R(A^*AB) \subseteq R(B)$ *and* $R(BB^*A^*) \subseteq R(A^*)$.

Proof. Let $N = AB$. Then by (2.1), (2.7) and (1.8), it follows that

$$\begin{aligned}
r(\, NN^\dagger - NB^\dagger A^\dagger \,) &= r\begin{bmatrix} N^*NN^* & N^* \\ NN^* & NB^\dagger A^\dagger \end{bmatrix} - r(N) \\
&= r\begin{bmatrix} 0 & N^* - N^*NB^\dagger A^\dagger \\ NN^* & 0 \end{bmatrix} - r(N) \\
&= r(\, N^* - N^*NB^\dagger A^\dagger \,) \\
&= r\begin{bmatrix} B^*A^* & B^*BB^* & 0 \\ A^*AA^* & 0 & A^* \\ 0 & N^*NB^* & -N^* \end{bmatrix} - r(A) - r(B) \\
&= r\begin{bmatrix} B^*A^* & B^*B & 0 \\ 0 & 0 & A^* \\ B^*A^*AA^* & N^*N & 0 \end{bmatrix} - r(A) - r(B) \\
&= r\begin{bmatrix} B^*A^* & B^*B \\ B^*A^*AA^* & N^*N \end{bmatrix} - r(B) \\
&= r\begin{bmatrix} B^*B & B^*A^*AB \\ AB & AA^*AB \end{bmatrix} - r(B) \\
&= r\left(\begin{bmatrix} B^* \\ A \end{bmatrix} [\, B, \ A^*AB \,] \right) - r(B) = r[\, B, \ A^*AB \,] - r(B),
\end{aligned}$$

as required for the first equality in Part (a). Applying (1.2) to it the block matrix in it yields the second equality in Part (a). Similarly, we can establish Part (b). The results in Parts (c) and (d) are direct consequences of Parts (a) and (b). The result in Part (e) follows directly from Parts (c) and (d). □

The result in Theorem 8.2(e) is well known, see, e.g., Arghiriade [4], Rao and Mitra [118], Ben-Israel and Greville [16], Campbell and Meyer [21]. Now it can be regarded as a direct consequence of some rank equalities related to Moore-Penrose inverses of products of two matrices. We next present another group rank equalities related to Moore-Penrose inverses of products of two matrices, which can also help to establish necessary and sufficient conditions for $(AB)^\dagger = B^\dagger A^\dagger$.

Theorem 8.3. *Let* $A \in \mathcal{C}^{m\times n}$ *and* $B \in \mathcal{C}^{n\times p}$ *be given. Then*

(a) $r[\, ABB^\dagger - (AB)(AB)^\dagger A \,] = r[\, B, \ A^*AB \,] - r(B)$.

(b) $r[\, A^\dagger AB - B(AB)^\dagger(AB) \,] = r\begin{bmatrix} A \\ ABB^* \end{bmatrix} - r(A)$.

(c) $r[\, A^*ABB^\dagger - BB^\dagger A^*A \,] = 2r[\, B, \ A^*AB \,] - 2r(B)$.

(d) $r[\, A^\dagger ABB^* - BB^*A^\dagger A \,] = 2r\begin{bmatrix} A \\ ABB^* \end{bmatrix} - 2r(A)$.

In particular,

(e) $(AB)(AB)^\dagger A = ABB^\dagger \Leftrightarrow A^*ABB^\dagger = BB^\dagger A^*A \Leftrightarrow R(A^*AB) \subseteq R(B) \Leftrightarrow B^\dagger A^\dagger \subseteq \{\, (AB)^{(1,2,3)} \}$.

(f) $A^\dagger AB = B(AB)^\dagger(AB) \Leftrightarrow A^\dagger ABB^* = BB^*A^\dagger A \Leftrightarrow R(BB^*A^*) \subseteq R(A^*) \Leftrightarrow B^\dagger A^\dagger \subseteq \{\, (AB)^{(1,2,4)} \}$.

(g) *The following three statements are equivalent (Greville [50])*:
(1) $(AB)^\dagger = B^\dagger A^\dagger$.
(2) $(AB)(AB)^\dagger A = ABB^\dagger$ *and* $A^\dagger AB = B(AB)^\dagger(AB)$.
(3) $A^*ABB^\dagger = BB^\dagger A^*A$ *and* $A^\dagger ABB^* = BB^*A^\dagger A$.

Proof. We only show Part (b). Note that both $AA^\dagger$ and $(AB)^\dagger(AB)$ are idempotent. We have by (3.1) that

$$r[\, A^\dagger AB - B(AB)^\dagger(AB) \,]$$

$$= r\begin{bmatrix} A^{\dagger}AB \\ (AB)^{\dagger}(AB) \end{bmatrix} + r[\, B(AB)^{\dagger}(AB),\ A^{\dagger}A\,] - r(A^{\dagger}A) - r[(AB)^{\dagger}(AB)]$$
$$= r(AB) + r[\, B(AB)^{*},\ A^{*}\,] - r(A) - r(AB) = r[\, BB^{*}A^{*},\ A^{*}\,] - r(A),$$

as required. □

Theorem 8.4. *Let $A \in \mathcal{C}^{m\times n}$ and $B \in \mathcal{C}^{n\times p}$ be given. Then*

(a) $r[\, A^{\dagger} - B(AB)^{\dagger}\,] = r\begin{bmatrix} A \\ ABB^{*} \end{bmatrix} - r(AB)$.

(b) $r[\, B^{\dagger} - (AB)^{\dagger}A\,] = r[\, B,\ A^{*}AB\,] - r(AB)$.

In particular,

(c) $A^{\dagger} = B(AB)^{\dagger} \Leftrightarrow R(A^{*}) = R(BB^{*}A^{*})$.

(d) $B^{\dagger} = (AB)^{\dagger}A \Leftrightarrow R(A^{*}) = R(A^{*}AB)$.

Proof. We only show Part (a). According to (2.2), we find that

$$\begin{aligned} r[\, A^{\dagger} - B(AB)^{\dagger}\,] &= r\begin{bmatrix} A^{*}AA^{*} & 0 & A^{*} \\ 0 & -(AB)^{*}AB(AB)^{*} & (AB)^{*} \\ A^{*} & B(AB)^{*} & 0 \end{bmatrix} - r(A) - r(AB) \\ &= r\begin{bmatrix} A^{*}AA^{*} & 0 & A^{*} \\ (AB)^{*}AA^{*} & & (AB)^{*} \\ A^{*} & B(AB)^{*} & 0 \end{bmatrix} - r(A) - r(AB) \\ &= r\begin{bmatrix} 0 & 0 & A^{*} \\ 0 & & (AB)^{*} \\ A^{*} & B(AB)^{*} & 0 \end{bmatrix} - r(A) - r(AB) \\ &= r\begin{bmatrix} A \\ ABB^{*} \end{bmatrix} - r(AB), \end{aligned}$$

establishing Part (a). □

We next consider ranks of matrix expressions involving Moore-Penrose inverses of products of three matrices, and then present their consequences related to reverse order laws.

Theorem 8.5. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times p}$ and $C \in \mathcal{C}^{p\times q}$ be given, and let $M = ABC$. Then*

$$r[\, M - M(BC)^{\dagger}B(AB)^{\dagger}M\,] = r\left(\begin{bmatrix} (BC)^{*} \\ A \end{bmatrix} B[\,(AB)^{*},\ C\,]\right) + r(M) - r(AB) - r(BC). \tag{8.2}$$

In particular,

$$(BC)^{\dagger}B(AB)^{\dagger} \in \{(ABC)^{-}\} \Leftrightarrow r\left(\begin{bmatrix} (BC)^{*} \\ A \end{bmatrix} B[\,(AB)^{*},\ C\,]\right) = r(AB) + r(BC) - r(M). \tag{8.3}$$

Proof. Applying (2.8) and the rank cancellation law (1.8) to $M - M(BC)^{\dagger}B(AB)^{\dagger}M$, we obtain

$$\begin{aligned} & r[\, M - M(BC)^{\dagger}B(AB)^{\dagger}M\,] \\ &= r\begin{bmatrix} (BC)^{*}B(AB)^{*} & (BC)^{*}(BC)(BC)^{*} & 0 \\ (AB)^{*}(AB)(AB)^{*} & 0 & (AB)^{*}M \\ 0 & M(BC)^{*} & -M \end{bmatrix} - r(AB) - r(BC) \\ &= r\begin{bmatrix} (BC)^{*}B(AB)^{*} & (BC)^{*}(BC) & 0 \\ (AB)(AB)^{*} & 0 & M \\ 0 & M & -M \end{bmatrix} - r(AB) - r(BC) \\ &= r\begin{bmatrix} (BC)^{*}B(AB)^{*} & (BC)^{*}(BC) \\ (AB)(AB)^{*} & M \end{bmatrix} + r(M) - r(AB) - r(BC) \\ &= r\left(\begin{bmatrix} (BC)^{*} \\ A \end{bmatrix} B[\,(AB)^{*},\ C\,]\right) + r(M) - r(AB) - r(BC). \end{aligned}$$

Thus we have (8.2) and (8.3). $\square$

As an application of (8.2), we consider the matrix product $M = (I_m + A)A(I_m - A) = A - A^3$. Then

$$\begin{aligned} & r[\,(A - A^3) - (A - A^3)(A - A^2)^{\dagger}A(A + A^2)^{\dagger}(A - A^3)\,] \\ & = r\left(\begin{bmatrix} (A - A^2)^* \\ I_m + A \end{bmatrix} A[\,(A + A^2)^*, \ I_m - A\,] \right) + r(A - A^3) - r(A + A^2) - r(A - A^2) \\ & = r\left(\begin{bmatrix} (A - A^2)^* \\ I_m + A \end{bmatrix} A[\,(A + A^2)^*, \ I_m - A\,] \right) - r(A) \leq 0. \end{aligned}$$

Notice that the rank of a matrix is nonnegative. The above inequality in fact implies that

$$r\left(\begin{bmatrix} (A - A^2)^* \\ I_m + A \end{bmatrix} A[\,(A + A^2)^*, \ I_m - A\,] \right) = r(A) \ \text{ and } \ (A-A^3)(A-A^2)^{\dagger}A(A+A^2)^{\dagger}(A-A^3) = (A-A^3),$$

that is, the matrix product $(A - A^2)^{\dagger}A(A + A^2)^{\dagger}$ is an inner inverse of the matrix $A - A^3$. By symmetry, $(A + A^2)^{\dagger}A(A - A^2)^{\dagger}$ is also an inner inverse of the matrix $A - A^3$.

In general, for any three polynomials $p_1(\lambda)$, $p_2(\lambda)$ and $p_3(\lambda)$ without common roots and a square matrix A, we let $p(A) = p_1(A)p_2(A)p_3(A)$. Then we can find by (8.2) and (1.17) the following

$$\begin{aligned} & r[\,p(A) - p(A)[p_2(A)p_3(A)]^{\dagger}p_2(A)[p_1(A)p_2(A)]^{\dagger}p(A)\,] \\ & = r\left(\begin{bmatrix} p_2(A^*)p_3(A^*) \\ p_1(A) \end{bmatrix} p_2(A)[\,[p_1(A^*)p_2(A^*), \ p_3(A)\,] \right) + r[p(A)] - r[p_1(A)p_2(A)] - r[p_2(A)p_3(A)] \\ & = r\left(\begin{bmatrix} p_2(A^*)p_3(A^*) \\ p_1(A) \end{bmatrix} p_2(A)[\,[p_1(A^*)p_2(A^*), \ p_3(A)\,] \right) - r[p_2(A)] \leq 0. \end{aligned}$$

This implies that

$$r\left(\begin{bmatrix} p_2(A^*)p_3(A^*) \\ p_1(A) \end{bmatrix} p_2(A)[\,[p_1(A^*)p_2(A^*), \ p_3(A)\,] \right) = r[p_2(A)]$$

and

$$p(A)[p_2(A)p_3(A)]^{\dagger}p_2(A)[p_1(A)p_2(A)]^{\dagger}p(A) = p(A).$$

Thus $[p_2(A)p_3(A)]^{\dagger}p_2(A)[p_1(A)p_2(A)]^{\dagger}$ is an inner inverse of the matrix product $p_1(A)p_2(A)p_3(A)$. By symmetry,

$$[p_1(A)p_2(A)]^{\dagger}p_2(A)[p_2(A)p_3(A)]^{\dagger}, \qquad [p_1(A)p_2(A)]^{\dagger}p_1(A)[p_1(A)p_3(A)]^{\dagger}$$

$$[p_1(A)p_3(A)]^{\dagger}p_1(A)[p_1(A)p_2(A)]^{\dagger}, \quad [p_1(A)p_3(A)]^{\dagger}p_3(A)[p_2(A)p_3(A)]^{\dagger}, \quad [p_2(A)p_3(A)]^{\dagger}p_3(A)[p_1(A)p_3(A)]^{\dagger}$$

are all inner inverses of the matrix product $p_1(A)p_2(A)p_3(A)$.

Theorem 8.6. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times p}$ and $C \in \mathcal{C}^{p\times q}$ be given, and let $M = ABC$. Then*

(a) *The rank of $M^{\dagger} - (BC)^{\dagger}B(AB)^{\dagger}$ satisfies the equality*

$$r[\,M^{\dagger} - (BC)^{\dagger}B(AB)^{\dagger}\,] = r\left(\begin{bmatrix} (BC)^* \\ M^*A \end{bmatrix} B[\,(AB)^*, \ CM^*\,] \right) - r(M). \tag{8.4}$$

(b) *The following three statements are equivalent*:

(1) $(ABC)^{\dagger} = (BC)^{\dagger}B(AB)^{\dagger}$.

(2) $r\begin{bmatrix} MM^*M & M(BC)^*(BC) \\ (AB)(AB)^*M & ABB^*BC \end{bmatrix} = r(ABC)$.

(3) $ABB^*BC = AB(BCM^{\dagger}AB)^*BC$.

(c) *If $r(ABC) = r(B)$, then*

$$(ABC)^{\dagger} = (BC)^{\dagger}B(AB)^{\dagger} \ \ \textit{and} \ \ (ABC)^{\dagger} = (B^{\dagger}BC)^{\dagger}B^{\dagger}(ABB^{\dagger})^{\dagger}. \tag{8.5}$$

Proof. Applying (2.12) to $M^\dagger - (BC)^\dagger B(AB)^\dagger$, we obtain

$$
\begin{aligned}
& r[\,M^\dagger - (BC)^\dagger B(AB)^\dagger\,] \\
= \ & r\begin{bmatrix} M^*MM^* & 0 & 0 & M^* \\ 0 & (BC)^*B(AB)^* & (BC)^*(BC)(BC)^* & 0 \\ 0 & (AB)^*(AB)(AB)^* & 0 & (AB)^* \\ M^* & 0 & (BC)^* & 0 \end{bmatrix} - r(M) - r(AB) - r(BC) \\
= \ & r\begin{bmatrix} M^*MM^* & -M^*(AB)(AB)^* & 0 & 0 \\ -(BC)^*(BC)M^* & (BC)^*B(AB)^* & 0 & 0 \\ 0 & 0 & 0 & (AB)^* \\ 0 & 0 & (BC)^* & 0 \end{bmatrix} - r(M) - r(AB) - r(BC) \\
= \ & r\begin{bmatrix} (BC)^*B(AB)^* & (BC)^*(BC)M^* \\ M^*(AB)(AB)^* & M^*MM^* \end{bmatrix} - r(M) \\
= \ & r\left(\begin{bmatrix} (BC)^* \\ M^*A \end{bmatrix} B[\,(AB)^*,\ CM^*\,] \right) - r(M),
\end{aligned}
$$

as required for (8.4). Then the equivalence of Statements (1) and (2) in Part (b) follows immediately from (8.4), and the equivalence of Statements (2) and (3) in Part (b) follows from Lemma 1.2(f). If $r(ABC) = r(B)$, then

$$
r\begin{bmatrix} MM^*M & M(BC)^*(BC) \\ (AB)(AB)^*M & ABB^*BC \end{bmatrix} \geq r(MM^*M) = r(M).
$$

On the other hand,

$$
r\begin{bmatrix} MM^*M & M(BC)^*(BC) \\ (AB)(AB)^*M & ABB^*BC \end{bmatrix} = r\left(\begin{bmatrix} MC^* \\ AB \end{bmatrix} B^*[\,A^*M,\ BC\,] \right) \leq r(B) = r(M).
$$

Thus we have

$$
r\begin{bmatrix} MM^*M & M(BC)^*(BC) \\ (AB)(AB)^*M & ABB^*BC \end{bmatrix} = r(M).
$$

Thus according to the statements (1) and (2) in Part (b), we know that the first equality in (8.5) is true. The second equality in (8.5) follows from writing $ABC = ABB^\dagger BC$ and then applying the first equality to it. $\square$

Theorem 8.7. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times p}$ and $C \in \mathcal{C}^{p\times q}$ be given, and let $M = ABC$. Then*

$$
r[\,B^\dagger - (AB)^\dagger M(BC)^\dagger\,] = r\begin{bmatrix} M & (AB)(AB)^* \\ (BC)^*(BC) & (BC)^*B(AB)^* \end{bmatrix} + r(B) - r(AB) - r(BC). \tag{8.6}
$$

In particular,

$$
B^\dagger = (AB)^\dagger M(BC)^\dagger \Leftrightarrow r\begin{bmatrix} M & (AB)(AB)^* \\ (BC)^*(BC) & (BC)^*B(AB)^* \end{bmatrix} = r(AB) + r(BC) - r(B).
$$

Proof. Applying (2.11) to $B^\dagger - (AB)^\dagger M(BC)^\dagger$, we obtain

$$
\begin{aligned}
& r[\,B^\dagger - (AB)^\dagger M(BC)^\dagger\,] \\
= \ & r\begin{bmatrix} B^*BB^* & 0 & 0 & B^* \\ 0 & (AB)^*M(BC)^* & (AB)^*(AB)(AB)^* & 0 \\ 0 & (BC)^*(BC)(BC)^* & 0 & (BC)^* \\ B^* & 0 & (AB)^* & 0 \end{bmatrix} - r(B) - r(AB) - r(BC) \\
= \ & r\begin{bmatrix} 0 & 0 & 0 & B^* \\ 0 & (AB)^*M(BC)^* & (AB)^*(AB)(AB)^* & 0 \\ 0 & (BC)^*(BC)(BC)^* & (BC)^*B(AB)^* & 0 \\ B^* & 0 & 0 & 0 \end{bmatrix} - r(B) - r(AB) - r(BC)
\end{aligned}
$$

$$
\begin{aligned}
&= r\begin{bmatrix} (AB)^*M(BC)^* & (AB)^*(AB)(AB)^* \\ (BC)^*(BC)(BC)^* & (BC)^*B(AB)^* \end{bmatrix} + r(B) - r(AB) - r(BC) \\
&= r\begin{bmatrix} M & (AB)(AB)^* \\ (BC)^*(BC) & (BC)^*B(AB)^* \end{bmatrix} + r(B) - r(AB) - r(BC),
\end{aligned}
$$

as required for (8.6). $\square$

Theorem 8.8. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times p}$ and $C \in \mathcal{C}^{p\times q}$ be given. Then*
(a) $(ABC)^\dagger = (A^\dagger ABC)^\dagger B(ABCC^\dagger)^\dagger$.
(b) $(ABC)^\dagger = [(AB)^\dagger ABC]^\dagger B^\dagger [ABC(BC)^\dagger]^\dagger$.
(c) (Cline [30]) $(AB)^\dagger = (A^\dagger AB)^\dagger (ABB^\dagger)^\dagger$.
(d) *If $R(C) \subseteq R[(AB)^*]$ and $R(A^*) \subseteq R(BC)$, then $(ABC)^\dagger = C^\dagger B^\dagger A^\dagger$.*

Proof. Write ABC as $ABC = A(A^\dagger ABCC^\dagger)C$. Then it is evident that

$$r(A^\dagger ABCC^\dagger) = r(ABC), \quad R[(ABCC^\dagger)^\dagger] \subseteq R(C), \quad \text{and} \quad R[((A^\dagger ABC)^\dagger)^*] \subseteq R(A^*).$$

Thus by (8.5), we find that

$$(ABC)^\dagger = [\,A(A^\dagger ABCC^\dagger)C\,]^\dagger = (A^\dagger ABC)^\dagger A^\dagger ABCC^\dagger (ABCC^\dagger)^\dagger = (A^\dagger ABC)^\dagger B(ABCC^\dagger)^\dagger,$$

as required for Part (a). On the other hand, we can write ABC as $ABC = (AB)B^\dagger(BC)$. Applying the equality in Part(a) to it yields

$$(ABC)^\dagger = [\,(AB)B^\dagger(BC)\,]^\dagger = [(AB)^\dagger ABC]^\dagger B^\dagger [ABC(BC)^\dagger]^\dagger,$$

as required for Part (b). Let B be identity matrix and replace C by B in the result in Part (a). Then we have the result in Part (c). The two conditions in Part (d) are equivalent to

$$(AB)^\dagger ABC = C, \quad \text{and} \quad ABC(BC)^\dagger = A.$$

In that case, the result in Part (b) reduces to the result in Part (d). $\square$

In the remainder of this chapter we consider the relationship of $(ABC)^\dagger$ and the reverse order product $C^\dagger B^\dagger A^\dagger$, and present necessary and sufficient conditions for $(ABC)^\dagger = C^\dagger B^\dagger A^\dagger$ to hold. Some of the results were presented by the first author in [133] and [135].

Lemma 8.9[135]. *Suppose that A_1, A_2, A_3, B_1 and B_2 satisfy the the following range inclusions*

$$R(B_i) \subseteq R(A_{i+1}), \quad \textit{and} \quad R(B_i^*) \subseteq R(A_i^*), \qquad i = 1,\ 2. \tag{8.7}$$

Then

$$\begin{bmatrix} 0 & 0 & A_1 \\ 0 & A_2 & B_1 \\ A_3 & B_2 & 0 \end{bmatrix}^\dagger = \begin{bmatrix} A_3^\dagger B_2 A_2^\dagger B_1 A_1^\dagger & -A_3^\dagger B_2 A_2^\dagger & A_3^\dagger \\ -A_2^\dagger B_1 A_1^\dagger & A_2^\dagger & 0 \\ A_1^\dagger & 0 & 0 \end{bmatrix}. \tag{8.8}$$

Proof. The range inclusions in (8.7) are equivalent to

$$A_{i+1}A_{i+1}^\dagger B_i = B_i, \quad \text{and} \quad B_i A_i^\dagger A_i = B_i, \qquad i = 1,\ 2.$$

In that case, it is easy to verify that the block matrix in the right-hand side of (8.8) and the given block matrix in the left-hand side of (8.8) satisfy the four Penrose equations. Thus (8.8) holds. $\square$

Lemma 8.10. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times p}$ and $C \in \mathcal{C}^{p\times q}$ be given. Then the product $C^\dagger B^\dagger A^\dagger$ can be written as*

$$C^\dagger B^\dagger A^\dagger = [\,I_q,\ 0,\ 0\,] \begin{bmatrix} 0 & 0 & AA^* \\ 0 & B^*BB^* & B^*A^* \\ C^*C & C^*B^* & 0 \end{bmatrix}^\dagger \begin{bmatrix} I_m \\ 0 \\ 0 \end{bmatrix} := PJ^\dagger Q, \tag{8.9}$$

where the block matrices P, J and Q satisfy

$$r(J) = r(A) + r(B) + r(C), \quad R(QA) \subseteq R(J), \quad \text{and} \quad R[(AP)^*] \subseteq R(J^*). \tag{8.10}$$

Proof. Observe that

$$R(B^*A^*) \subseteq R(B^*BB^*), \quad R(AB) \subseteq R(AA^*), \quad R(C^*B^*) \subseteq R(C^*C), \quad R(BC) \subseteq R(BB^*B),$$

as well as the three basic equalities on the Moore-Penrose inverse of a matrix

$$N^\dagger = N^*(N^*NN^*)^\dagger N^*, \qquad N^\dagger = (N^*N)^\dagger N^*, \qquad N^\dagger = N^*(NN^*)^\dagger.$$

Thus we find by (8.8) that

$$\begin{aligned} \begin{bmatrix} 0 & 0 & AA^* \\ 0 & B^*BB^* & B^*A \\ C^*C & C^*B^* & 0 \end{bmatrix}^\dagger &= \begin{bmatrix} (C^*C)^\dagger C^*B^*(BB^*B)^\dagger B^*A^*(AA^*)^\dagger & * & * \\ * & * & 0 \\ * & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} C^\dagger B^\dagger A^\dagger & * & * \\ * & * & 0 \\ * & 0 & 0 \end{bmatrix}. \end{aligned}$$

Hence we have (8.9). The properties in (8.10) are obvious. $\square$

Theorem 8.11. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times p}$ and $C \in \mathcal{C}^{p\times q}$ be given and let $M = ABC$. Then*

$$r[\, M - M(C^\dagger B^\dagger A^\dagger)M \,] = r\begin{bmatrix} -M^* & 0 & C^*C \\ 0 & BB^*B & BC \\ AA^* & AB & 0 \end{bmatrix} - r(A) - r(B) - r(C) + r(M). \tag{8.11}$$

In particular,

$$C^\dagger B^\dagger A^\dagger \in \{(ABC)^-\} \Leftrightarrow r\begin{bmatrix} -M^* & 0 & C^*C \\ 0 & BB^*B & BC \\ AA^* & AB & 0 \end{bmatrix} = r(A) + r(B) + r(C) - r(M). \tag{8.12}$$

Proof. It follows from (1.7) and (8.9) that

$$\begin{aligned} r[\, M - M(C^\dagger B^\dagger A^\dagger)M \,] &= r(\, M - MPJ^\dagger QM \,) \\ &= r\begin{bmatrix} J & QM \\ MP & M \end{bmatrix} - r(J) \\ &= r\begin{bmatrix} J - QMP & 0 \\ 0 & M \end{bmatrix} - r(J) \\ &= r(\, J - QMP \,) + r(M) - r(J) \\ &= r\begin{bmatrix} -M & 0 & AA^* \\ 0 & B^*BB^* & B^*A^* \\ C^*C & C^*B^* & 0 \end{bmatrix} + r(M) - r(J) \\ &= r\begin{bmatrix} -M^* & 0 & C^*C \\ 0 & BB^*B & BC \\ AA^* & AB & 0 \end{bmatrix} + r(M) - r(A) - r(B) - r(C), \end{aligned}$$

as required for (8.11). $\square$

Applying (8.11) to the matrix product $M = (I_m + A)A(I_m - A) = A - A^3$, we can find that

$$r[\, M - M(I_m - A)^\dagger A^\dagger (I_m + A)^\dagger M \,] = 0.$$

Thus $(I_m - A)^\dagger A^\dagger (I_m + A)^\dagger$ is an inner inverse of the matrix $A - A^3$. By symmetry, $(I_m - A)^\dagger A^\dagger (I_m + A)^\dagger$ is also an inner inverse of the matrix $A - A^3$. We leave the verification of the rank equality to the reader.

Applying (8.11) to the matrix product $M = (I_m + A)A(I_m - A) = A - A^3$, we can find that

$$r[\, M - M(I_m - A)^{\dagger}A^{\dagger}(I_m + A)^{\dagger}M \,] = 0.$$

Thus $(I_m - A)^{\dagger}A^{\dagger}(I_m + A)^{\dagger}$ is an inner inverse of the matrix $A - A^3$. By symmetry, $(I_m - A)^{\dagger}A^{\dagger}(I_m + A)^{\dagger}$ is also an inner inverse of the matrix $A - A^3$. We leave the verification of the rank equality to the reader.

In general, for any three polynomials $p_1(\lambda)$, $p_2(\lambda)$ and $p_3(\lambda)$ without common roots and a square matrix A, we let $p(A) = p_1(A)p_2(A)p_3(A)$. Then one can find by (8.11) and (1.17) the following

$$r[\, p(A) - p(A)p_3^{\dagger}(A)p_2^{\dagger}(A)p_1^{\dagger}(A)p(A) \,] = 0.$$

Thus $p_3^{\dagger}(A)p_2^{\dagger}(A)p_1^{\dagger}(A)$ is an inner inverse inverse of the matrix product $p_1(A)p_2(A)p_3(A)$.

Theorem 8.12. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times p}$ and $C \in \mathcal{C}^{p\times q}$ be given and let $M = ABC$. Then*

$$r(\, M^{\dagger} - C^{\dagger}B^{\dagger}A^{\dagger} \,) = r\begin{bmatrix} -MM^*M & 0 & MC^*C \\ 0 & BB^*B & BC \\ AA^*M & AB & 0 \end{bmatrix} - r(B) - r(M). \tag{8.13}$$

In particular,

$$(ABC)^{\dagger} = C^{\dagger}B^{\dagger}A^{\dagger} \Leftrightarrow r\begin{bmatrix} -MM^*M & 0 & MC^*C \\ 0 & BB^*B & BC \\ AA^*M & AB & 0 \end{bmatrix} = r(B) + r(ABC). \tag{8.14}$$

Proof. Notice that

$$C^{\dagger}CM^{\dagger}AA^{\dagger} = M^{\dagger} \quad \text{and} \quad C^{\dagger}C(C^{\dagger}B^{\dagger}A^{\dagger})AA^{\dagger} = C^{\dagger}B^{\dagger}A^{\dagger}.$$

We first get the following

$$\begin{aligned} r(\, M^{\dagger} - C^{\dagger}B^{\dagger}A^{\dagger} \,) &= r(\, CM^{\dagger}A - CC^{\dagger}B^{\dagger}A^{\dagger}A \,) \\ &= r(\, CM^{\dagger}A - CPJ^{\dagger}QA \,) \\ &= r[\, CM^*(M^*MM^*)^{\dagger}M^*A - CPJ^{\dagger}QA \,] \\ &= r\left([\, CM^*,\ CP \,] \begin{bmatrix} -M^*MM^* & 0 \\ 0 & J \end{bmatrix}^{\dagger} \begin{bmatrix} M^*A \\ QA \end{bmatrix} \right). \end{aligned}$$

Observe from (8.10) that

$$R\begin{bmatrix} M^*A \\ QA \end{bmatrix} \subseteq R\begin{bmatrix} -M^*MM^* & 0 \\ 0 & J \end{bmatrix} \quad \text{and} \quad R(\, [\, CM^*,\ CP \,]^* \,) \subseteq R\left(\begin{bmatrix} -M^*MM^* & 0 \\ 0 & J \end{bmatrix}^* \right).$$

Thus we find by (1.7) that

$$\begin{aligned} & r\left([\, CM^*,\ CP \,] \begin{bmatrix} -M^*MM^* & 0 \\ 0 & J \end{bmatrix}^{\dagger} \begin{bmatrix} M^*A \\ QA \end{bmatrix} \right) \\ &= r\begin{bmatrix} -MM^*M & 0 & M^*A \\ 0 & J & QA \\ CM^* & CP & 0 \end{bmatrix} - r\begin{bmatrix} -M^*MM^* & 0 \\ 0 & J \end{bmatrix} \\ &= r\begin{bmatrix} -MM^*M & 0 & 0 & 0 & M^*A \\ 0 & 0 & 0 & AA^* & A \\ 0 & 0 & B^*BB^* & B^*A^* & 0 \\ 0 & C^*C & C^*B^* & 0 & 0 \\ CM^* & C & 0 & 0 & 0 \end{bmatrix} - r(M) - r(J) \\ &= r\begin{bmatrix} -MM^*M & 0 & 0 & 0 - M^*AA^* & 0 \\ 0 & 0 & 0 & 0 & A \\ 0 & 0 & B^*BB^* & B^*A^* & 0 \\ -C^*CM^* & 0 & C^*B^* & 0 & 0 \\ 0 & C & 0 & 0 & 0 \end{bmatrix} - r(M) - r(J) \end{aligned}$$

$$= r\begin{bmatrix} -MM^*M & 0 & -M^*AA^* \\ 0 & B^*BB^* & B^*A^* \\ -C^*CM^* & C^*B^* & 0 \end{bmatrix} - r(M) - r(B).$$

The results in (8.13) and (8.14) follow from it. □

Corollary 8.13. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times p}$ and $C \in \mathcal{C}^{p\times q}$ be given, and let $M = ABC$. If*

$$R(B) \subseteq R(A^*) \quad and \quad R(B^*) \subseteq R(C), \tag{8.15}$$

then

$$r(\, M^\dagger - C^\dagger B^\dagger A^\dagger \,) = r\begin{bmatrix} B \\ BCC^* \end{bmatrix} + r[\, B, \; A^*AB \,] - 2r(B). \tag{8.16}$$

In particular,

$$(ABC)^\dagger = C^\dagger B^\dagger A^\dagger \Leftrightarrow R(A^*AB) \subseteq R(B) \quad and \quad R[(BCC^*)^*] \subseteq R(B^*). \tag{8.17}$$

Proof. Eq. (8.15) is equivalent to $A^\dagger AB = B$ and $BCC^\dagger = B$. Thus we can reduce (8.13) by block elementary operations to

$$\begin{aligned} r(\, M^\dagger - C^\dagger B^\dagger A^\dagger \,) &= r\begin{bmatrix} -BCM^*AB & 0 & BCC^* \\ 0 & BB^*B & B \\ A^*AB & B & 0 \end{bmatrix} - 2r(B) \\ &= r\begin{bmatrix} 0 & BCC^*B^* & BCC^* \\ 0 & BB^*B & B \\ A^*AB & B & 0 \end{bmatrix} - 2r(B) \\ &= r\begin{bmatrix} 0 & 0 & BCC^*B \\ 0 & 0 & B \\ A^*AB & B & 0 \end{bmatrix} - 2r(B) \\ &= r\begin{bmatrix} B \\ BCC^* \end{bmatrix} + r[\, B, \; A^*AB \,] - 2r(B). \qquad \square \end{aligned}$$

Corollary 8.14. *Let $B \in \mathcal{C}^{m\times n}$ be given, $A \in \mathcal{C}^{m\times m}$ and $C \in \mathcal{C}^{n\times n}$ be two invertible matrices. Let $M = ABC$. Then*

$$r[\, (ABC)^\dagger - C^{-1}B^\dagger A^{-1} \,] = r\begin{bmatrix} B \\ BCC^* \end{bmatrix} + r[\, B, \; A^*AB \,] - 2r(B), \tag{8.18}$$

and

$$r[\, (ABC)^\dagger - C^{-1}B^\dagger A^{-1} \,] = r\begin{bmatrix} M \\ MC^*C \end{bmatrix} + r[\, M, \; AA^*M \,] - 2r(M). \tag{8.19}$$

In particular,

$$(ABC)^\dagger = C^{-1}B^\dagger A^{-1} \Leftrightarrow R(AA^*B) = R(B) \quad and \quad R(CC^*B^*) = R(B^*), \tag{8.20}$$

and

$$(ABC)^\dagger = C^{-1}B^\dagger A^{-1} \Leftrightarrow R(AA^*M) = R(M) \quad and \quad R(C^*CM^*) = R(M^*). \tag{8.21}$$

Proof. Follows immediately from Corollary 8.13. □

Theorem 8.15. *Let $B \in \mathcal{C}^{m\times n}$ be given, $A \in \mathcal{C}^{m\times m}$ and $C \in \mathcal{C}^{n\times n}$ be two invertible matrices. Let $M = ABC$. Then*

(a) $r(\, MM^\dagger - ABB^\dagger A^{-1} \,) = r[\, B, \; A^*AB \,] - r(B).$

(b) $r(\, M^\dagger M - C^{-1}B^\dagger BC \,) = r\begin{bmatrix} B \\ BCC^* \end{bmatrix} - r(B).$

(c) $r(\, M^\dagger - C^{-1}B^\dagger A^{-1} \,) = r(\, MM^\dagger - ABB^\dagger A^{-1} \,) + r(\, M^\dagger M - C^{-1}B^\dagger BC \,).$

In particular,

(d) $MM^\dagger = ABB^\dagger A^{-1} \Leftrightarrow R(A^*AB) = R(B).$

(e) $M^\dagger M = C^{-1}B^\dagger BC \Leftrightarrow R(CC^*B^*) = R(B^*).$

(f) $M^\dagger = C^{-1}B^\dagger A^{-1} \Leftrightarrow MM^\dagger = ABB^\dagger A^{-1} \Leftrightarrow M^\dagger M = C^{-1}B^\dagger BC.$

Proof. Observe that

$$r(\,MM^{\dagger} - ABB^{\dagger}A^{-1}\,) = r(\,MM^{\dagger}A - ABB^{\dagger}\,) \;\text{ and }\; r(\,M^{\dagger}M - C^{-1}B^{\dagger}BC\,) = r(\,CM^{\dagger}M - B^{\dagger}BC\,).$$

Applying (7.4) to both of them yields Parts (a) and (b). Contrasting (8.18) with Parts (a) and (b) yields Part (c). □

Finally we present author interesting result on the Moore-Penrose inverse of a triple matrix product.

Corollary 8.16. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times p}$ and $C \in \mathcal{C}^{p\times q}$ be given, and suppose that $R(B) \subseteq R(A^*)$ and $R(B^*) \subseteq R(C)$. Then*

$$r[\,(ABC)^{\dagger} - (\,I_q - (C^{\dagger}F_B)(C^{\dagger}F_B)^{\dagger}\,)C^{\dagger}B^{\dagger}A^{\dagger}(\,I_m - (E_BA^{\dagger})^{\dagger}(E_BA^{\dagger})\,)\,] = m + q - r(A) - r(C), \tag{8.22}$$

where $E_B = I_n - BB^{\dagger}$ and $F_B = I_p - B^{\dagger}B$. In particular, the equality

$$(ABC)^{\dagger} = [\,I_q - (C^{\dagger}F_B)(C^{\dagger}F_B)^{\dagger}\,]C^{\dagger}B^{\dagger}A^{\dagger}[\,I_m - (E_BA^{\dagger})^{\dagger}(E_BA^{\dagger})\,] \tag{8.23}$$

holds if and only if $r(A) = m$ and $r(C) = q$. Thus if both A and C are nonsingular matrices, then the identity

$$(ABC)^{\dagger} = [\,I_q - (C^{-1}F_B)(C^{-1}F_B)^{\dagger}\,]C^{-1}B^{\dagger}A^{-1}[\,I_m - (E_BA^{-1})^{\dagger}(E_BA^{-1})\,] \tag{8.24}$$

holds.

Proof. Let $M = ABC$ and $N = [\,I_q - (C^{\dagger}F_B)(C^{\dagger}F_B)^{\dagger}\,]C^{\dagger}B^{\dagger}A^{\dagger}[\,I_m - (E_BA^{\dagger})^{\dagger}(E_BA^{\dagger})\,]$. Then it is easy to verify that under $R(B) \subseteq R(A^*)$ and $R(B^*) \subseteq R(C)$, N is an outer inverse of M. Hence by (5.1) we get

$$r(\,M^{\dagger} - N\,) = r\begin{bmatrix} M^{\dagger} \\ N \end{bmatrix} + r[\,M^{\dagger},\; N\,] - r(M) - r(N). \tag{8.25}$$

Simplifying the ranks of the matrices in (8.25) by (8.23) and (1.2)—(1.4), we can eventually get

$$r\begin{bmatrix} M^{\dagger} \\ N \end{bmatrix} = m + r(B) - r(C), \quad r[\,M^{\dagger},\; N\,] = q + r(A) - r(C), \quad r(M) = r(B), \quad r(N) = r(B).$$

The tedious processes are omitted here. Putting them in (8.25) we have (8.22), and then (8.23) and (8.24). □

It is expected that the identity (8.24) can help to establish various equalities for Moore-Penrose inverses of block matrices.

Chapter 9

Moore-Penrose inverses of block matrices

In this chapter we establish some rank equalities related to factorizations of 2×2 block matrices and then deduce from them various expressions of Moore-Penrose inverses for 2×2 block matrices, as well as for $m \times n$ block matrices. Some of the results in this chapter appear in the author's recent paper [136]. In fact, any 2×2 block matrix can simply factor as various types of products of block matrices. In that case, applying the rank equalities in Chapter 8 to them one can establish many new rank equalities related to the block matrix, and consequently, derive from them various expressions for the Moore-Penrose inverse of the block matrix. We begin this work first with a bordered matrix.

Theorem 9.1. *Let* $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$ *be a given bordered matrix over the field of complex numbers, where* $A \in \mathcal{C}^{m \times n}$, $B \in \mathcal{C}^{m \times k}$ *and* $C \in \mathcal{C}^{l \times n}$, *and factor* M *as*

$$M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = \begin{bmatrix} I_m & E_B A C^\dagger \\ 0 & I_l \end{bmatrix} \begin{bmatrix} E_B A F_C & B \\ C & 0 \end{bmatrix} \begin{bmatrix} I_n & 0 \\ B^\dagger A & I_k \end{bmatrix} := PNQ, \tag{9.1}$$

where $E_B = I_m - BB^\dagger$ *and* $F_C = I_n - C^\dagger C$. *Then*

(a) *The rank of* $M^\dagger - Q^{-1} N^\dagger P^{-1}$ *satisfies the equality*

$$r(M^\dagger - Q^{-1} N^\dagger P^{-1}) = r \begin{bmatrix} A \\ C \end{bmatrix} + r[\, A, \ B \,] + r(B) + r(C) - 2r(M). \tag{9.2}$$

(b) *The following four statements are equivalent*:

(1) *The Moore-Penrose inverse of* M *can be expressed as* $M^\dagger = Q^{-1} N^\dagger P^{-1}$, *that is,*

$$M^\dagger = \begin{bmatrix} (E_B A F_C)^\dagger & C^\dagger - (E_B A F_C)^\dagger A C^\dagger \\ B^\dagger - B^\dagger A (E_B A F_C)^\dagger & -B^\dagger A C^\dagger + B^\dagger A (E_B A F_C)^\dagger A C^\dagger \end{bmatrix}. \tag{9.3}$$

(2) $[\, I_n, \ 0 \,] M^\dagger \begin{bmatrix} I_m \\ 0 \end{bmatrix} = (E_B A F_C)^\dagger$.

(3) A, B *and* C *satisfy the rank additivity condition*

$$r(M) = r \begin{bmatrix} A \\ C \end{bmatrix} + r(B) = r[\, A, \ B \,] + r(C). \tag{9.4}$$

(4) *The two conditions hold*

$$R \begin{bmatrix} A \\ C \end{bmatrix} \cap R \begin{bmatrix} B \\ 0 \end{bmatrix} = \{0\} \ \ \text{and} \ \ R([\, A, \ B \,]^*) \cap R([\, C, \ 0 \,]^*) = \{0\}. \tag{9.5}$$

Proof. It follows first by (9.1) and (8.19) that

$$r(M^\dagger - Q^{-1} N^\dagger P^{-1}) = r \begin{bmatrix} M \\ MQ^*Q \end{bmatrix} + r[\, M, \ PP^*M \,] - 2r(M). \tag{9.6}$$

The ranks of the two block matrices in (9.6) can simplify to

$$\begin{aligned} r[M,\ PP^*M] &= r\left[\begin{bmatrix} A & B \\ C & 0 \end{bmatrix},\ \begin{bmatrix} I_m & E_BAC^\dagger \\ 0 & I_l \end{bmatrix}\begin{bmatrix} I_m & 0 \\ (E_BAC^\dagger)^* & I_l \end{bmatrix}\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\right] \\ &= r\begin{bmatrix} A & B & A+(E_BAC^\dagger)(E_BAC^\dagger)^*A + E_BAC^\dagger C & B \\ C & 0 & C+(E_BAC^\dagger)^*A & 0 \end{bmatrix} \\ &= r\begin{bmatrix} A & B & AC^\dagger(E_BAC^\dagger)^*A + AC^\dagger C \\ C & 0 & (E_BAC^\dagger)^*A \end{bmatrix} \\ &= r\begin{bmatrix} A & B & AC^\dagger C \\ C & 0 & 0 \end{bmatrix} \\ &= r\begin{bmatrix} A & B & 0 \\ C & 0 & -C \end{bmatrix} = r[A,\ B] + r(C), \end{aligned}$$

and

$$\begin{aligned} r\begin{bmatrix} M \\ MQ^*Q \end{bmatrix} &= r\begin{bmatrix} \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} \\ \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\begin{bmatrix} I_n & (B^\dagger A)^* \\ 0 & I_k \end{bmatrix}\begin{bmatrix} I_n & 0 \\ B^\dagger A & I_k \end{bmatrix} \end{bmatrix} \\ &= r\begin{bmatrix} A & B \\ C & 0 \\ A + A(B^\dagger A)^*(B^\dagger A) + BB^\dagger A & B + A(B^\dagger A)^* \\ C + C(B^\dagger A)^*B^\dagger A & C(B^\dagger A)^* \end{bmatrix} \\ &= r\begin{bmatrix} A & B \\ C & 0 \\ A + A(B^\dagger A)^*(B^\dagger A) & A(B^\dagger A)^* \\ C(B^\dagger A)^*B^\dagger A & C(B^\dagger A)^* \end{bmatrix} \\ &= r\begin{bmatrix} A & B \\ C & 0 \\ A & 0 \\ 0 & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r(B). \end{aligned}$$

Putting both of them in (9.6) yields (9.2). Notice that $(E_BAF_C)^\dagger B^\dagger = 0$ and $C^\dagger(E_BAF_C)^\dagger = 0$ always hold. Then it is easy to verify that

$$N^\dagger = \begin{bmatrix} E_BAF_C & B \\ C & 0 \end{bmatrix}^\dagger = \begin{bmatrix} (E_BAF_C)^\dagger & C^\dagger \\ B^\dagger & 0 \end{bmatrix}.$$

Putting it in (9.2), we get

$$Q^{-1}N^\dagger P^{-1} = \begin{bmatrix} (E_BAF_C)^\dagger & C^\dagger - (E_BAF_C)^\dagger AC^\dagger \\ B^\dagger - B^\dagger A(E_BAF_C)^\dagger & -B^\dagger AC^\dagger + B^\dagger A(E_BAF_C)^\dagger AC^\dagger \end{bmatrix}.$$

The equivalence of the statements (1) and (3) in Part (b) follows from (9.2). The equivalence of the statements (2) and (3) in Part (b) comes from (7.29). The equivalence of the statements (3) and (4) in Part (b) is obvious. □

The expression (9.3) for $M^\dagger$ is well known when M satisfies (9.4) (see, i.e. [100] and [118]). The rank equality (9.2) further reveals a fact that (9.3) is not only sufficient but also necessary. Various consequences can be derived from Theorem 9.1 when the matrix M in it satisfies some more restrictions. Here we only present one that is well known.

Corollary 9.2. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given. If $R(A)\cap R(B) = \{0\}$ and $R(A^*)\cap R(C^*) = \{0\}$, then*

$$\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^\dagger = \begin{bmatrix} (E_BAF_C)^\dagger & C^\dagger - (E_BAF_C)^\dagger AC^\dagger \\ B^\dagger - B^\dagger A(E_BAF_C)^\dagger & 0 \end{bmatrix}. \tag{9.7}$$

Proof. Under $R(A) \cap R(B) = \{0\}$ and $R(A^*) \cap R(C^*) = \{0\}$, the rank equality (9.4) naturally holds. In that case, we know by Theorem 7.8 that $A(E_B A F_C)^{\dagger} A = A$. Thus (9.3) reduces to (9.7). □

Besides the factorization (9.1), we can generally factor M as

$$M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = \begin{bmatrix} I_m & Y \\ 0 & I_l \end{bmatrix} \begin{bmatrix} A - BX - YC & B \\ C & 0 \end{bmatrix} \begin{bmatrix} I_n & 0 \\ X & I_k \end{bmatrix} := PNQ, \tag{9.8}$$

where X and Y are arbitrary and P and Q are nonsingular. Clearly (9.1) is a special case of (9.8). In that case, applying the first equality in (8.5) to (9.8) we obtain the following.

Theorem 9.3. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$ be given. Then*

$$\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^{\dagger} = \begin{bmatrix} A - YC & B \\ C & 0 \end{bmatrix}^{\dagger} \begin{bmatrix} A - BX - YC & B \\ C & 0 \end{bmatrix} \begin{bmatrix} A - BX & B \\ C & 0 \end{bmatrix}^{\dagger}, \tag{9.9}$$

where X and Y are arbitrary. In particular,

$$\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^{\dagger} = \left[[\, AF_C, \ B\,]^{\dagger}, \ \begin{bmatrix} C^{\dagger} \\ 0 \end{bmatrix} \right] \begin{bmatrix} A - BB^{\dagger}A - AC^{\dagger}C & B \\ C & 0 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} E_B A \\ C \end{bmatrix}^{\dagger} \\ [\, B^{\dagger}, \ 0\,] \end{bmatrix}. \tag{9.10}$$

Proof. Under (9.8) we have by (8.5) that $M = (PNQ)^{\dagger} = (NQ)^{\dagger} N (PN)^{\dagger}$. Written in an explicit form, it is (9.9). Now let $X = B^{\dagger}A$ and $Y = AC^{\dagger}$ in (9.9). Then (9.9) becomes

$$\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^{\dagger} = \begin{bmatrix} AF_C & B \\ C & 0 \end{bmatrix}^{\dagger} \begin{bmatrix} A - BB^{\dagger}A - AC^{\dagger}C & B \\ C & 0 \end{bmatrix} \begin{bmatrix} E_B A & B \\ C & 0 \end{bmatrix}^{\dagger}.$$

Note that

$$[\, AF_C, \ B\,][\, C, \ 0\,]^* = 0, \quad \text{and} \quad \begin{bmatrix} B \\ 0 \end{bmatrix}^* \begin{bmatrix} E_B A \\ C \end{bmatrix} = 0.$$

Then it follows by Theorem 7.9(e) and (f) that

$$\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^{\dagger} = [\, [\, AF_C, \ B\,]^{\dagger}, \ [\, C, \ 0\,]^{\dagger}\,], \quad \text{and} \quad \begin{bmatrix} E_B A & B \\ C & 0 \end{bmatrix}^{\dagger} = \begin{bmatrix} \begin{bmatrix} E_B A \\ C \end{bmatrix}^{\dagger} \\ \begin{bmatrix} B \\ 0 \end{bmatrix}^{\dagger} \end{bmatrix}.$$

Thus we have (9.10). □

Eq. (9.10) manifests that the Moore-Penrose inverse of a bordered matrix can be intermediately determined by the Moore-Penrose inverse of B, C, $[\, AF_C, \ B\,]$ and $\begin{bmatrix} E_B A \\ C \end{bmatrix}$. Observe that

$$[\, AF_C, \ B\,]^{\dagger} = [\, AF_C, \ B\,]^* ([\, AF_C, \ B\,][\, AF_C, \ B\,]^*)^{\dagger} = \begin{bmatrix} (AF_C)^* [\, (AF_C)(AF_C)^* + BB^*\,]^{\dagger} \\ B^* [\, (AF_C)(AF_C)^* + BB^*\,]^{\dagger} \end{bmatrix},$$

$$\begin{aligned} \begin{bmatrix} E_B A \\ C \end{bmatrix}^{\dagger} &= \left(\begin{bmatrix} E_B A \\ C \end{bmatrix}^* \begin{bmatrix} E_B A \\ C \end{bmatrix} \right)^{\dagger} \begin{bmatrix} E_B A \\ C \end{bmatrix}^* \\ &= [\, (\, (E_B A)^*(E_B A) + C^*C\,)^{\dagger}(E_B A)^*, \ (\, (E_B A)^*(E_B A) + C^*C\,)^{\dagger} C^*\,]. \end{aligned}$$

Inserting them in (9.10) we get

$$\begin{aligned} \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^{\dagger} = &\begin{bmatrix} (AF_C)^* [\, (AF_C)(AF_C)^* + BB^*\,]^{\dagger} & C^{\dagger} \\ B^* [\, (AF_C)(AF_C)^* + BB^*\,]^{\dagger} & 0 \end{bmatrix} \begin{bmatrix} A - BB^{\dagger}A - AC^{\dagger}C & B \\ C & 0 \end{bmatrix} \\ &\times \begin{bmatrix} [\, (E_B A)^*(E_B A) + C^*C\,]^{\dagger} \, (E_B A)^* & [\, (E_B A)^*(E_B A) + C^*C\,]^{\dagger} \, C^* \\ B^{\dagger} & 0 \end{bmatrix}, \end{aligned}$$

which could be regarded as a general expression for the Moore-Penrose inverse a boredered matrix when no restriction is posed on it. Moreover, this expression reveals another interesting that the Moore-Penrose inverse a boredered matrix can factor as a product of three boredered matrices although the Moore-Penrose inverse of the boredered matrix is not boredered in general.

Another well-known factorization for a bordered matrix is

$$M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ CA^\dagger & I_l \end{bmatrix} \begin{bmatrix} A & E_AB \\ CF_A & -CA^\dagger B \end{bmatrix} \begin{bmatrix} I_n & A^\dagger B \\ 0 & I_k \end{bmatrix}. \tag{9.11}$$

But it can be considered as a special case of the Schur factorization of a 2×2 block matrix,

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ CA^\dagger & I_l \end{bmatrix} \begin{bmatrix} A & E_AB \\ CF_A & S_A \end{bmatrix} \begin{bmatrix} I_n & A^\dagger B \\ 0 & I_k \end{bmatrix} := PNQ, \tag{9.12}$$

where $S_A = D - CA^\dagger B$. We next present a rank equality related to (9.12) and derive its consequences.

Theorem 9.4. *Let M be given by* (9.12), *where $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$, $C \in \mathcal{C}^{l\times n}$ and $D \in \mathcal{C}^{l\times k}$. Then the rank of $M^\dagger - Q^{-1}N^\dagger P^{-1}$ satisfies the equality*

$$r(\, M^\dagger - Q^{-1}N^\dagger P^{-1}\,) = r\begin{bmatrix} A & 0 \\ 0 & C \\ B & D \end{bmatrix} + r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} - 2r(M). \tag{9.13}$$

In particular, the Moore-Penrose inverse of M in (9.12) *can be expressed as $M^\dagger = Q^{-1}N^\dagger P^{-1}$, that is,*

$$M^\dagger = \begin{bmatrix} I_n & -A^\dagger B \\ 0 & I_k \end{bmatrix} \begin{bmatrix} A & E_AB \\ CF_A & S_A \end{bmatrix}^\dagger \begin{bmatrix} I_m & 0 \\ -CA^\dagger & I_l \end{bmatrix} \tag{9.14}$$

holds if and only if A, B, C and D satisfy

$$r\begin{bmatrix} A & 0 \\ 0 & C \\ B & D \end{bmatrix} = r(M) \quad and \quad r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} = r(M), \tag{9.15}$$

or equivalently

$$R\begin{bmatrix} A \\ 0 \end{bmatrix} \subseteq R\begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad and \quad R\begin{bmatrix} A^* \\ 0 \end{bmatrix} \subseteq R\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix}. \tag{9.16}$$

Proof. It follows by (9.12) and (8.19) that

$$r(\, M^\dagger - Q^{-1}N^\dagger P^{-1}\,) = r\begin{bmatrix} M \\ MQ^*Q \end{bmatrix} + r[\, M, \ PP^*M\,] - 2r(M).$$

The ranks of the two block matrices in it can reduce to

$$\begin{aligned} & r[\, M, \ PP^*M\,] \\ = \ & r\left[\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \begin{bmatrix} I & 0 \\ CA^\dagger & I \end{bmatrix} \begin{bmatrix} I & (CA^\dagger)^* \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \right] \\ = \ & r\begin{bmatrix} A & B & A + (CA^\dagger)^*C & B + (CA^\dagger)^*D \\ C & D & CA^\dagger A + (CA^\dagger)(CA^\dagger)^*C + C & CA^\dagger B + (CA^\dagger)(CA^\dagger)^*D + D \end{bmatrix} \\ = \ & r\begin{bmatrix} A & B & (CA^\dagger)^*C & (CA^\dagger)^*D \\ C & D & (CA^\dagger)(CA^\dagger)^*C + C & CA^\dagger B + (CA^\dagger)(CA^\dagger)^*D \end{bmatrix} \\ = \ & r\begin{bmatrix} A & B & 0 & 0 \\ C & D & C & CA^\dagger B \end{bmatrix} \\ = \ & r\begin{bmatrix} A & B & 0 \\ C & D & C \end{bmatrix} = r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix}. \end{aligned}$$

Similarly we can get

$$r\begin{bmatrix} M \\ MQ^*Q \end{bmatrix} = r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix}.$$

Thus we have (9.13). Eqs. (9.14)—(9.16) are direct consequences of (9.12). □

When $D = 0$ in Theorem 9.4, we the following.

Corollary 9.5. *Let M be given by (9.11) where $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$ and $C \in \mathcal{C}^{l\times n}$. Then the rank of $M^\dagger - Q^{-1}N^\dagger P^{-1}$ satisfies the equality*

$$r[\, M^\dagger - Q^{-1}N^\dagger P^{-1}\,] = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A, \ B\,] + r(B) + r(C) - 2r(M). \tag{9.17}$$

In particular, the Moore-Penrose inverse of M can be expressed as

$$M^\dagger = Q^{-1}N^\dagger P^{-1} = \begin{bmatrix} I_n & -A^\dagger B \\ 0 & I_k \end{bmatrix}\begin{bmatrix} A & E_AB \\ CF_A & -CA^\dagger B \end{bmatrix}^\dagger \begin{bmatrix} I_m & 0 \\ -CA^\dagger & I_l \end{bmatrix}, \tag{9.18}$$

if and only if $A, \ B$ and C satisfy the following rank additivity condition

$$r(M) = r\begin{bmatrix} A \\ C \end{bmatrix} + r(B) = r[\, A, \ B\,] + r(C). \tag{9.19}$$

Eq. (9.19) shows that we have another expression for the Moore-Penrose inverse of a bordered matrix M when it satisfies the rank additivity condition (9.19) (the first one is in (9.3)).

Clearly the matrix N in (9.12) can be written as

$$N = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & E_AB \\ CF_A & S_A \end{bmatrix} = N_1 + N_2. \tag{9.20}$$

Then it is easy to verify that

$$N^\dagger = \begin{bmatrix} A^\dagger & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & E_AB \\ CF_A & S_A \end{bmatrix}^\dagger = N_1^\dagger + N_2^\dagger. \tag{9.21}$$

Thus if we can find $N_2^\dagger$, then we can give the expression of $N^\dagger$ in (9.21). This consideration motivates us to find the following set of results on Moore-Penrose inverses of block matrices.

Lemma 9.6. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$, $C \in \mathcal{C}^{l\times n}$ and $D \in \mathcal{C}^{l\times k}$ be given. Then the rank additivity condition*

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r\begin{bmatrix} B \\ D \end{bmatrix} = r[\, A, \ B\,] + r[\, C, \ D\,] \tag{9.22}$$

is equivalent to the two range inclusions

$$R\begin{bmatrix} A \\ 0 \end{bmatrix} \subseteq R\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \qquad R\begin{bmatrix} A^* \\ 0 \end{bmatrix} \subseteq R\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix}, \tag{9.23}$$

and the rank additivity condition

$$r\begin{bmatrix} 0 & E_AB \\ CF_A & S_A \end{bmatrix} = r\begin{bmatrix} E_AB \\ S_A \end{bmatrix} + r(CF_A) = r[\, CF_A, \ S_A\,] + r(E_AB), \tag{9.24}$$

where $S_A = D - CA^\dagger B$.

Proof. Let

$$V_1 = \begin{bmatrix} A \\ C \end{bmatrix}, \quad V_2 = \begin{bmatrix} B \\ D \end{bmatrix}, \quad W_1 = [\,A,\ B\,], \quad W_2 = [\,C,\ D\,]. \tag{9.25}$$

Then (9.22) is equivalent to

$$R(V_1) \cap R(V_2) = \{0\} \qquad \text{and} \qquad R(W_1^*) \cap R(W_2^*) = \{0\}. \tag{9.26}$$

In that case, we easily find

$$r\begin{bmatrix} W_1 & A \\ W_2 & 0 \end{bmatrix} = r[\,W_1,\ A\,] + r[\,W_2,\ 0\,] = r(W_1) + r(W_2) = r(M),$$

and

$$r\begin{bmatrix} V_1 & V_2 \\ A & 0 \end{bmatrix} = r\begin{bmatrix} V_1 \\ A \end{bmatrix} + r\begin{bmatrix} V_2 \\ 0 \end{bmatrix} = r(V_1) + r(V_2) = r(M),$$

both of which are equivalent to the two inclusions in (9.23). On the other hand, observe that

$$\begin{bmatrix} E_AB \\ S_A \end{bmatrix} = \begin{bmatrix} B - AA^\dagger B \\ D - CA^\dagger B \end{bmatrix} = \begin{bmatrix} B \\ D \end{bmatrix} - \begin{bmatrix} A \\ C \end{bmatrix} A^\dagger B = V_2 - V_1A^\dagger B,$$

and

$$[\,CF_A,\ S_A\,] = [\,C - CAA^\dagger,\ D - CA^\dagger B\,] = [\,C,\ D\,] - CA^\dagger[\,A,\ B\,] = W_2 - CA^\dagger W_1.$$

Thus according to (9.26) and Lemma 1.4(b) and (c), we find that

$$r\begin{bmatrix} E_AB \\ S_A \end{bmatrix} = r(\,V_2 - V_1A^\dagger B\,) = r\begin{bmatrix} V_2 \\ V_1A^\dagger B \end{bmatrix} = r(V_2), \tag{9.27}$$

and

$$r[\,CF_A,\ S_A\,] = r(\,W_2 - CA^\dagger W_1\,) = r[\,W_2,\ CA^\dagger W_1\,] = r(W_2). \tag{9.28}$$

From both of them and the rank formulas in (1.2), (1.3), (1.5), (9.23), (9.27) and (9.28), we derive the following two equalities

$$r\begin{bmatrix} 0 & E_AB \\ CF_A & S_A \end{bmatrix} = r(M) - r(A) = r(V_1) + r(V_2) - r(A) = r\begin{bmatrix} E_AB \\ S_A \end{bmatrix} + r(CF_A),$$

and

$$r\begin{bmatrix} 0 & E_AB \\ CF_A & S_A \end{bmatrix} = r(M) - r(A) = r(W_1) + r(W_2) - r(A) = r[\,CF_A,\ S_A\,] + r(E_AB).$$

Both of them are exactly the rank additivity condition (9.24). Conversely, adding $r(A)$ to the three sides of (9.24) and then applying (1.2), (1.3) and (1.5) to the corresponding result we first obtain

$$r\begin{bmatrix} A & E_AB \\ CF_A & S_A \end{bmatrix} = r\begin{bmatrix} A \\ CF_A \end{bmatrix} + r\begin{bmatrix} E_AB \\ S_A \end{bmatrix} = r[\,A,\ E_AB\,] + r[\,CF_A,\ S_A\,]. \tag{9.29}$$

On the other hand, the two inclusions in (9.23) are also equivalent to

$$r(M) = r\begin{bmatrix} A & B & A \\ C & D & 0 \end{bmatrix} = r\begin{bmatrix} A & B & 0 \\ C & D & C \end{bmatrix}, \ \text{ and } \ r(M) = r\begin{bmatrix} A & B \\ C & D \\ A & 0 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \\ 0 & B \end{bmatrix}.$$

Applying (1.5) to the right-hand sides of the above two equalities and then combining them with (9.29), we find

$$r(M) = r\begin{bmatrix} A & E_AB & 0 \\ CF_A & S_A & C \end{bmatrix} = r[\,A,\ E_AB\,] + r[\,CF_A,\ S_A,\ C\,] = r[\,A,\ B\,] + r[\,C,\ D\,],$$

and

$$r(M) = r\begin{bmatrix} A & E_AB \\ CF_A & S_A \\ 0 & B \end{bmatrix} = r\begin{bmatrix} A \\ CF_A \end{bmatrix} + r\begin{bmatrix} E_AB \\ S_A \\ B \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r\begin{bmatrix} B \\ D \end{bmatrix}.$$

Both of them are exactly (9.22). $\square$

Similarly we can establish the following.

Lemma 9.7. *The rank additivity condition* (9.22) *is equivalent to the following four conditions*

$$R\begin{bmatrix} A \\ 0 \end{bmatrix} \subseteq R\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \qquad R\begin{bmatrix} A^* \\ 0 \end{bmatrix} \subseteq R\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix}, \tag{9.30}$$

and

$$r\begin{bmatrix} S_D & BF_D \\ E_DC & 0 \end{bmatrix} = r\begin{bmatrix} S_D \\ E_DC \end{bmatrix} + r(BF_D) = r[\, S_D,\ BF_D \,] + r(E_DC), \tag{9.31}$$

where $S_D = A - BD^\dagger C$.

Theorem 9.8. *Suppose that the block matrix M in* (9.11) *satisfies the rank additivity condition* (9.22), *then the Moore-Penrose inverse of M can be expressed in the two forms*

$$M^\dagger = \begin{bmatrix} H_1 - H_2CA^\dagger - A^\dagger BH_3 + A^\dagger BJ^\dagger(D)CA^\dagger & H_2 - A^\dagger BJ^\dagger(D) \\ H_3 - J^\dagger(D)CA^\dagger & J^\dagger(D) \end{bmatrix}, \tag{9.32}$$

and

$$M^\dagger = \begin{bmatrix} J^\dagger(A) & J^\dagger(C) \\ J^\dagger(B) & J^\dagger(D) \end{bmatrix} = \begin{bmatrix} (E_{B_2}S_DF_{C_2})^\dagger & (E_{D_2}S_BF_{A_1})^\dagger \\ (E_{A_2}S_CF_{D_1})^\dagger & (E_{C_1}S_AF_{B_1})^\dagger \end{bmatrix}, \tag{9.33}$$

where

$$\begin{gathered} S_A = D - CA^\dagger B, \quad S_B = C - DB^\dagger A, \quad S_C = B - AC^\dagger D, \quad S_D = A - BD^\dagger C, \\ A_1 = E_BA, \qquad A_2 = AF_C, \qquad B_1 = E_AB, \quad B_2 = BF_D, \\ C_1 = CF_A, \qquad C_2 = E_DC, \qquad D_1 = E_CD, \quad D_2 = DF_B, \\ H_1 = A^\dagger + C_1^\dagger[\, S_AJ^\dagger(D)S_A - S_A \,]B_1^\dagger, \quad H_2 = C_1^\dagger[\, I - S_AJ^\dagger(D) \,], \quad H_3 = [\, I - J^\dagger(D)S_A \,]B_1^\dagger. \end{gathered}$$

Proof. Lemma 9.6 shows that the rank additivity condition in (9.22) is equivalent to (9.23) and (9.24). It follows from Theorem 9.4 that under (9.23), the Moore-Penrose inverse of M can be expressed as (9.13). On the other hand, It follows from Theorem 9.1 that under (9.24) the Moore-Penrose inverse of N_2 in (9.20) can be written as

$$N_2^\dagger = \begin{bmatrix} C_1^\dagger[\, S_AJ^\dagger(D)S_A - S_A \,]B_1^\dagger & C_1^\dagger - C_1^\dagger S_AJ^\dagger(D) \\ B_1^\dagger - J^\dagger(D)S_AB_1^\dagger & J^\dagger(D) \end{bmatrix}, \tag{9.34}$$

where $B_1 = E_AB$, $C_1 = CF_A$ and $J(D) = E_{C_1}S_AF_{B_1}$. Now substituting (9.34) into (9.21) and then (9.21) into (9.13), we get

$$M^\dagger = Q^{-1}N^\dagger P^{-1} = Q^{-1}\begin{bmatrix} A^\dagger + C_1^\dagger[\, S_AJ^\dagger(D)S_A - S_A \,]B_1^\dagger & C_1^\dagger - C_1^\dagger S_AJ^\dagger(D) \\ B_1^\dagger - J^\dagger(D)S_AB_1^\dagger & J^\dagger(D) \end{bmatrix}P^{-1}. \tag{9.35}$$

Written in a 2×2 block matrix, (9.35) is (9.32). In the same way, we can also decompose M in (9.11) into the other three forms,

$$M = \begin{bmatrix} I_m & 0 \\ CB^\dagger & I_l \end{bmatrix}\begin{bmatrix} E_BA & B \\ S_B & DF_B \end{bmatrix}\begin{bmatrix} I_n & 0 \\ B^\dagger A & I_k \end{bmatrix},$$

$$M = \begin{bmatrix} I_m & AC^\dagger \\ 0 & I_l \end{bmatrix}\begin{bmatrix} AF_C & S_C \\ C & E_CD \end{bmatrix}\begin{bmatrix} I_n & C^\dagger D \\ 0 & I_k \end{bmatrix},$$

and

$$M = \begin{bmatrix} I_m & BD^\dagger \\ 0 & I_l \end{bmatrix}\begin{bmatrix} S_D & BF_D \\ E_DC & D \end{bmatrix}\begin{bmatrix} I_n & 0 \\ D^\dagger C & I_k \end{bmatrix}.$$

Based on the above decompositions of M we can also find that under (9.22) the Moore-Penrose inverse of M can also be expressed as

$$M^\dagger = \begin{bmatrix} * & J^\dagger(C) \\ * & * \end{bmatrix} = \begin{bmatrix} * & * \\ J^\dagger(B) & * \end{bmatrix} = \begin{bmatrix} J^\dagger(A) & * \\ * & * \end{bmatrix}. \tag{9.36}$$

Finally from the uniqueness of the Moore-Penrose inverse of a matrix and the expressions in (9.32) and (9.36), we obtain (9.33). □

Some fundamental properties on the Moore-Penrose inverse of M in (9.11) can be derive from (9.32) and (9.33).

Corollary 9.9. *Denote the Moore-Penrose inverse of M in Eq. (9.11) by*

$$M^{\dagger} = \begin{bmatrix} G_1 & G_2 \\ G_3 & G_4 \end{bmatrix}, \tag{9.37}$$

where G_1, G_2, G_3 and G_4 are $n \times m$, $n \times l$, $k \times m$ and $k \times l$ matrices, respectively. If M in (9.11) satisfies the rank additivity condition (9.22), then the submatrices in M and $M^{\dagger}$ satisfy the rank equalities

$$r(G_1) = r(V_1) + r(W_1) - r(M) + r(D), \tag{9.38}$$

$$r(G_2) = r(V_1) + r(W_2) - r(M) + r(B), \tag{9.39}$$

$$r(G_3) = r(V_2) + r(W_1) - r(M) + r(C), \tag{9.40}$$

$$r(G_4) = r(V_2) + r(W_2) - r(M) + r(A), \tag{9.41}$$

$$r(G_1) + r(G_4) = r(A) + r(D), \qquad r(G_2) + r(G_3) = r(B) + r(C), \tag{9.42}$$

where V_1, V_2, W_1 and W_2 are defined in (9.25). Moreover, the products of $MM^{\dagger}$ and $M^{\dagger}M$ have the forms

$$MM^{\dagger} = \begin{bmatrix} W_1W_1^{\dagger} & 0 \\ 0 & W_2W_2^{\dagger} \end{bmatrix}, \qquad M^{\dagger}M = \begin{bmatrix} V_1^{\dagger}V_1 & 0 \\ 0 & V_2^{\dagger}V_2 \end{bmatrix}. \tag{9.43}$$

Proof. The four rank equalities in (9.38)—(9.42) can directly be derived from the expression in (9.33) for $M^{\dagger}$ and the rank formula (1.6). The two equalities in (9.42) come from the sums of (9.38) and (9.41), (9.39) and (9.40), respectively. The two results in (9.43) are derived from (9.22) and Theorem 7.16(c) and (d). □

The rank additivity condition (9.22) is a quite weak restriction to a 2×2 block matrix. As a matter of fact, any matrix with its rank great then 1 satisfies a rank additivity condition as in (9.22) when its rows and columns are properly permuted. We next present a group of consequences of Theorem 9.8.

Corollary 9.10. *If the block matrix M in (9.11) satisfies (9.23) and the following two conditions*

$$R(C_1) \cap R(S_A) = \{0\} \quad \text{and} \quad R(B_1^*) \cap R(S_A^*) = \{0\}, \tag{9.44}$$

then the Moore-Penrose inverse of M can be expressed as

$$\begin{aligned} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger} &= Q^{-1} \begin{bmatrix} A^{\dagger} & C_1^{\dagger} - C_1^{\dagger}S_AJ^{\dagger}(D) \\ B_1^{\dagger} - J^{\dagger}(D)S_AB_1^{\dagger} & J^{\dagger}(D) \end{bmatrix} P^{-1} \\ &= \begin{bmatrix} A^{\dagger} - H_2CA^{\dagger} - A^{\dagger}BH_3 + A^{\dagger}BJ^{\dagger}(D)CA^{\dagger} & H_2 - A^{\dagger}BJ^{\dagger}(D) \\ H_3 - J^{\dagger}(D)CA^{\dagger} & J^{\dagger}(D) \end{bmatrix}, \end{aligned}$$

where C_1, B_1, H_2, H_3 and $J(D)$ are as in (9.32), P and Q are as in (9.13).

Proof. The conditions in (9.44) imply that the block matrix N_2 in (9.20) satisfies the following rank additivity condition

$$r(N_2) = r(E_AB) + r(CF_A) + r(S_A),$$

which is a special case of (9.24). On the other hand, under (9.44) if follow by Theorem 7.7 that $S_AJ^{\dagger}(D)S_A = S_A$. Thus (9.35) reduces to the desired result in the corollary. □

Corollary 9.11. *If the block matrix M in (9.11) satisfies (9.23) and the two conditions*

$$R(BS_A^*) \subseteq R(A) \quad \text{and} \quad R(C^*S_A) \subseteq R(A^*), \tag{9.45}$$

then the Moore-Penrose inverse of M can be expressed as

$$\begin{aligned} M^{\dagger} &= \begin{bmatrix} I_n & -A^{\dagger}B \\ 0 & I_k \end{bmatrix} \begin{bmatrix} A^{\dagger} & (CF_A)^{\dagger} \\ (E_AB)^{\dagger} & S_A^{\dagger} \end{bmatrix} \begin{bmatrix} I_m & 0 \\ -CA^{\dagger} & I_l \end{bmatrix} \\ &= \begin{bmatrix} A^{\dagger} - A^{\dagger}B(E_AB)^{\dagger} - (CF_A)^{\dagger}CA^{\dagger} + A^{\dagger}BS_A^{\dagger}CA^{\dagger} & (CF_A)^{\dagger} - A^{\dagger}BS_A^{\dagger} \\ (E_AB)^{\dagger} - S_A^{\dagger}CA^{\dagger} & S_A^{\dagger} \end{bmatrix}, \end{aligned}$$

where $S_A = D - CA^{\dagger}B$.

Proof. Clearly (9.45) are equivalent to $(E_AB)S_A^* = 0$ and $S_A^*(CF_A) = 0$. In that case, (9.24) is satisfied, and $N^{\dagger} = \begin{bmatrix} A^{\dagger} & (CF_A)^{\dagger} \\ (E_AB)^{\dagger} & S_A^{\dagger} \end{bmatrix}$ in (9.35). □

Corollary 9.12(Chen and Zhou [27]). *If the block matrix M in* (9.11) *satisfies the following four conditions*

$$R(B) \subseteq R(A), \qquad R(C^*) \subseteq R(A^*), \qquad R(C) \subseteq R(S_A), \qquad R(B^*) \subseteq R(S_A^*), \tag{9.46}$$

then the Moore-Penrose inverse of M can be expressed as

$$\begin{aligned} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger} &= \begin{bmatrix} I_n & -A^{\dagger}B \\ 0 & I_k \end{bmatrix} \begin{bmatrix} A^{\dagger} & 0 \\ 0 & S_A^{\dagger} \end{bmatrix} \begin{bmatrix} I_m & 0 \\ -CA^{\dagger} & I_l \end{bmatrix} \\ &= \begin{bmatrix} A^{\dagger} + A^{\dagger}BS_A^{\dagger}CA^{\dagger} & -A^{\dagger}BS_A^{\dagger} \\ -S_A^{\dagger}CA^{\dagger} & S_A^{\dagger} \end{bmatrix}, \end{aligned}$$

where $S_A = D - CA^{\dagger}B$.

Proof. It is easy to verify that under the conditions in (9.46), the rank of M satisfies the rank additivity condition (9.22). In that case, $N^{\dagger} = \begin{bmatrix} A^{\dagger} & 0 \\ 0 & S_A^{\dagger} \end{bmatrix}$ in (9.35). □

Corollary 9.13. *If the block matrix M in* (9.11) *satisfies the four conditions*

$$R(A) \cap R(B) = \{0\}, \quad R(A^*) \cap R(C^*) = \{0\}, \quad R(D) \subseteq R(C), \quad R(D^*) \subseteq R(B^*), \tag{9.47}$$

then the Moore-Penrose inverse of M can be expressed as

$$\begin{aligned} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger} &= \begin{bmatrix} I_n & -A^{\dagger}B \\ 0 & I_k \end{bmatrix} \begin{bmatrix} A^{\dagger} - C_1^{\dagger}S_AB_1^{\dagger} & C_1^{\dagger} \\ B_1^{\dagger} & 0 \end{bmatrix} \begin{bmatrix} I_m & 0 \\ -CA^{\dagger} & I_l \end{bmatrix} \\ &= \begin{bmatrix} A^{\dagger} - A^{\dagger}BB_1^{\dagger} - C_1^{\dagger}CA^{\dagger} - C_1^{\dagger}S_AB_1^{\dagger} & C_1^{\dagger} \\ B_1^{\dagger} & 0 \end{bmatrix}, \end{aligned}$$

where $S_A = D - CA^{\dagger}B$, $B_1 = E_AB$ and $C_1 = CF_A$.

Proof. It is not difficult to verify by (1.5) that under (9.47) the rank of M satisfies (9.22). In that case, $J(D) = 0$ and $N^{\dagger} = \begin{bmatrix} A^{\dagger} - C_1^{\dagger}S_AB_1^{\dagger} & C_1^{\dagger} \\ B_1^{\dagger} & 0 \end{bmatrix}$ in (9.35). □

Corollary 9.14. *If the block matrix M in* (9.11) *satisfies the four conditions*

$$R(A) \cap R(B) = \{0\}, \qquad R(A^*) \cap R(C^*) = \{0\}, \tag{9.48}$$

$$R(S_A) \subseteq N(C^*), \qquad R(S_A^*) \subseteq N(B), \tag{9.49}$$

then the Moore-Penrose inverse of M can be expressed as

$$\begin{aligned} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger} &= \begin{bmatrix} I_n & -A^{\dagger}B \\ 0 & I_k \end{bmatrix} \begin{bmatrix} A^{\dagger} & (CF_A)^{\dagger} \\ (E_AB)^{\dagger} & S_A^{\dagger} \end{bmatrix} \begin{bmatrix} I_m & 0 \\ -CA^{\dagger} & I_l \end{bmatrix} \\ &= \begin{bmatrix} A^{\dagger} - A^{\dagger}B(E_AB)^{\dagger} - (CF_A)^{\dagger}CA^{\dagger} & (CF_A)^{\dagger} \\ (E_AB)^{\dagger} & S_A^{\dagger} \end{bmatrix}, \end{aligned}$$

where $S_A = D - CA^{\dagger}B$.

Proof. Clearly (9.49) is equivalent to $C^{\dagger}S_A = 0$ and $S_AB^{\dagger} = 0$, as well as $S_A^{\dagger}C = 0$ and $BS_A^{\dagger} = 0$. From them and (9.48), we also find

$$(CF_A)^{\dagger}S_A = 0 \quad \text{and} \quad S_A(E_AB)^{\dagger} = 0. \tag{9.50}$$

Combining (9.48) and (9.50) shows that M satisfies (9.23) and (9.24). In that case,

$$N^{\dagger} = \begin{bmatrix} A^{\dagger} & (CF_A)^{\dagger} \\ (E_AB)^{\dagger} & S_A^{\dagger} \end{bmatrix}$$

in (9.35). □

Corollary 9.15. *If the block matrix* M *in* (9.11) *satisfies the rank additivity condition*

$$r(M_1) = r(A) + r(B) + r(C) + r(D), \tag{9.51}$$

then the Moore-Penrose inverse of M *can be expressed as*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger} = \begin{bmatrix} (E_BAF_C)^{\dagger} & (E_DCF_A)^{\dagger} \\ (E_ABF_D)^{\dagger} & (E_CDF_B)^{\dagger} \end{bmatrix}. \tag{9.52}$$

Proof. Obviously (9.51) is a special case of (9.22). On the other hand, (9.51) is also equivalent to the following four conditions

$$R(A) \cap R(B) = \{0\}, \quad R(C) \cap R(D) = \{0\}, \quad R(A^*) \cap R(C^*) = \{0\}, \quad R(B^*) \cap R(D) = \{0\}.$$

In that case,

$$\begin{array}{llll} R(A_1^*) = R(A^*), & R(A_2) = R(A), & R(A_1^*) = R(A^*), & R(B_2) = R(B), \\ R(C_1) = R(A), & R(C_2^*) = R(C^*), & R(D_1^*) = R(D^*), & R(A_2) = R(D), \end{array}$$

by Lemma 1.2(a) and (b). Then it turns out by Theorem 7.2(c) and (d) that

$$\begin{array}{llll} A_1^{\dagger}A_1 = A^{\dagger}A, & A_2A_2^{\dagger} = AA^{\dagger}, & B_1^{\dagger}B_1 = B^{\dagger}B, & B_2B_2^{\dagger} = BB^{\dagger}, \\ C_1C_1^{\dagger} = CC^{\dagger}, & C_2^{\dagger}C_2 = C^{\dagger}C, & D_1^{\dagger}D_1 = D^{\dagger}D, & D_2D_2^{\dagger} = DD^{\dagger}. \end{array}$$

Thus (9.33) reduces to (9.52) □

Corollary 9.16. *If the block matrix* M *in Eq.* (9.11) *satisfies* $r(M) = r(A) + r(D)$ *and*

$$R(B) \subseteq R(A), \quad R(C) \subseteq R(D), \quad R(C^*) \subseteq R(A^*), \quad R(B^*) \subseteq R(D^*),$$

then the Moore-Penrose inverse of M *can be expressed as*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger} = \begin{bmatrix} (A - BD^{\dagger}C)^{\dagger} & -A^{\dagger}B(D - CA^{\dagger}B)^{\dagger} \\ -(D - CA^{\dagger}B)^{\dagger}CA^{\dagger} & (D - CA^{\dagger}B)^{\dagger} \end{bmatrix}.$$

Corollary 9.17. *If the block matrix* M *in* (9.11) *satisfies* $r(M) = r(A) + r(D)$ *and the following four conditions*

$$R(A) = R(B), \quad R(C) = R(D), \quad R(A^*) = R(C^*), \quad R(B^*) = R(D^*),$$

then the Moore-Penrose inverse of M *can be expressed as*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger} = \begin{bmatrix} S_D^{\dagger} & S_B^{\dagger} \\ S_C^{\dagger} & S_A^{\dagger} \end{bmatrix} = \begin{bmatrix} (A - BD^{\dagger}C)^{\dagger} & (C - DB^{\dagger}A)^{\dagger} \\ (B - AC^{\dagger}D)^{\dagger} & (D - CA^{\dagger}B)^{\dagger} \end{bmatrix}.$$

The above two corollaries can directly be derived from (9.32) and (9.33), the proofs are omitted here.

Without much effort, we can extend the results in Theorem 9.8 to $m \times n$ block matrices when they satisfy rank additivity conditions.

Let

$$M = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix} \tag{9.53}$$

be an $m \times n$ block matrix , where A_{ij} is an $s_i \times t_j$ matrix $(1 \le i \le m,\ 1 \le j \le n)$, and suppose that M satisfies the following rank additivity condition

$$r(M) = r(W_1) + r(W_2) + \cdots + r(W_m) = r(V_1) + r(V_2) + \cdots + r(V_n), \tag{9.54}$$

where

$$W_i = [\, A_{i1},\, A_{i2},\, \cdots,\, A_{in}\,], \qquad V_j = \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{bmatrix}, \qquad 1 \le i \le m,\ 1 \le j \le n. \tag{9.55}$$

For convenience of representation, we adopt the following notation. Let $M = (A_{ij})$ be given in (9.53), where $A_{ij} \in \mathcal{C}^{s_i \times t_j}$, $1 \le i \le m$, $1 \le j \le n$, and $\sum_{i=1}^m s_i = s$, $\sum_{i=1}^n t_i = t$. For each A_{ij} in M we associate three block matrices as follows

$$B_{ij} = [\, A_{i1},\, \cdots,\, A_{i,j-1},\, A_{i,j+1},\, \cdots,\, A_{in}\,], \tag{9.56}$$

$$C_{ij}^* = [\, A_{1j}^*,\, \cdots,\, ,A_{i-1,j}^*,\, A_{i+1,j}^*,\, \cdots,\, A_{mj}^*\,], \tag{9.57}$$

$$D_{ij} = \begin{bmatrix} A_{11} & \cdots & A_{1,j-1} & A_{1,j+1} & \cdots & A_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{i-1,1} & \cdots & A_{i-1,j-1} & A_{i-1,j+1} & \cdots & A_{i-1,n} \\ A_{i+1,1} & \cdots & A_{i+1,j-1} & A_{i+1,j+1} & \cdots & A_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{m1} & \cdots & A_{m,j-1} & A_{m,j+1} & \cdots & A_{mn} \end{bmatrix}. \tag{9.58}$$

The symbol $J(A_{ij})$ stands for

$$J(A_{ij}) = E_{\alpha_{ij}} S_{D_{ij}} F_{\beta_{ij}}, \qquad 1 \le i \le m,\ 1 \le j \le n, \tag{9.59}$$

where $\alpha_{ij} = B_{ij}F_{D_{ij}}$, $\beta_{ij} = E_{D_{ij}}C_{ij}$ and $S_{D_{ij}} = A_{ij} - B_{ij}D_{ij}^{\dagger}C_{ij}$ is the Schur complement of D_{ij} in M. We call the matrix $J(A_{ij})$ the *rank complement* of A_{ij} in M. Besides we partition the Moore-Penrose inverse of M in (9.53) into the form

$$M^{\dagger} = \begin{bmatrix} G_{11} & G_{12} & \cdots & G_{1m} \\ G_{21} & G_{22} & \cdots & G_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ G_{n1} & G_{n2} & \cdots & G_{nm} \end{bmatrix}, \tag{9.60}$$

where G_{ij} is a $t_i \times s_j$ matrix, $1 \le i \le n$, $1 \le j \le m$.

Next we build two groups of block permutation matrices as follows

$$P_1 = I_s, \qquad P_i = \begin{bmatrix} 0 & & & I_{s_i} & & & \\ I_{s_1} & \ddots & & & & & \\ & \ddots & \ddots & & & & \\ & & I_{s_{i-1}} & 0 & & & \\ & & & & I_{s_{i+1}} & & \\ & & & & & \ddots & \\ & & & & & & I_{s_m} \end{bmatrix}, \tag{9.61}$$

$$Q_1 = I_s, \qquad Q_j = \begin{bmatrix} 0 & I_{t_1} & & & & & & \\ & \ddots & \ddots & & & & & \\ & & \ddots & I_{t_{j-1}} & & & & \\ I_{t_j} & & & 0 & & & & \\ & & & & I_{t_{j+1}} & & & \\ & & & & & \ddots & \\ & & & & & & I_{t_n} \end{bmatrix}, \tag{9.62}$$

where $2 \le i \le m,\ 2 \le j \le n$. Applying (9.61) and (9.62) to M in (9.53) and $M^\dagger$ in (9.60) we have the following two groups of expressions

$$P_iMQ_j = \begin{bmatrix} A_{ij} & B_{ij} \\ C_{ij} & D_{ij} \end{bmatrix}, \qquad 1 \le i \le m,\ 1 \le j \le n, \tag{9.63}$$

and

$$Q_j^T M^\dagger P_i^T = \begin{bmatrix} G_{ji} & * \\ * & * \end{bmatrix}, \qquad 1 \le i \le m,\ 1 \le j \le n. \tag{9.64}$$

These two equalities show that we can use two block permutation matrices to permute A_{ij} in M and the corresponding block G_{ji} in $M^\dagger$ to the upper left corners of M and $M^\dagger$, respectively. Observe that P_i and Q_j in (9.61) and (9.62) are all orthogonal matrices. The Moore-Penrose inverse of P_iMQ_j in (9.63) can be expressed as $(P_iMQ_j)^\dagger = Q_j^T M^\dagger P_i^T$. Combining (9.63) with (9.64), we have the following simple result

$$\begin{bmatrix} A_{ij} & B_{ij} \\ C_{ij} & D_{ij} \end{bmatrix}^\dagger = \begin{bmatrix} G_{ji} & * \\ * & * \end{bmatrix}, \qquad 1 \le i \le m,\ 1 \le j \le n. \tag{9.65}$$

If the block matrix M in (9.53) satisfies the rank additivity condition (9.54), then the 2×2 block matrix on the right-hand side of (9.63) naturally satisfies the following rank additivity condition

$$r\begin{bmatrix} A_{ij} & B_{ij} \\ C_{ij} & D_{ij} \end{bmatrix} = r\begin{bmatrix} A_{ij} \\ C_{ij} \end{bmatrix} + r\begin{bmatrix} B_{ij} \\ D_{ij} \end{bmatrix} = r[\, A_{ij},\ B_{ij}\,] + r[\, C_{ij},\ D_{ij}\,], \tag{9.66}$$

where $1 \le i \le m,\ 1 \le j \le n$. Hence combining (9.65) and (9.66) with Theorems 9.8 and 9.9, we obtain the following general result.

Theorem 9.18. *Suppose that the $m \times n$ block matrix M in* (9.53) *satisfies the rank additivity condition* (9.54). *Then*

(a) *The Moore-Penrose inverse of M can be expressed as*

$$M^\dagger = \begin{bmatrix} J^\dagger(A_{11}) & J^\dagger(A_{21}) & \cdots & J^\dagger(A_{m1}) \\ J^\dagger(A_{12}) & J^\dagger(A_{22}) & \cdots & J^\dagger(A_{m2}) \\ \cdots & \cdots & \cdots & \cdots \\ J^\dagger(A_{1n}) & J^\dagger(A_{2n}) & \cdots & J^\dagger(A_{mn}) \end{bmatrix}, \tag{9.67}$$

where $J(A_{ij})$ is defined in (9.59).

(b) *The rank of the block entry $G_{ji} = J^\dagger(A_{ij})$ in $M^\dagger$ is*

$$r(G_{ji}) = r[J(A_{ij})] = r(W_i) + r(V_j) - r(M) + r(D_{ij}), \tag{9.68}$$

where $1 \le i \le m,\ 1 \le j \le n$, W_i and V_j are defined in (9.55).

(c) *$MM^\dagger$ and $M^\dagger M$ are two block diagonal matrices*

$$MM^\dagger = \mathrm{diag}(\, W_1W_1^\dagger,\ W_2W_2^\dagger,\ \cdots,\ W_mW_m^\dagger\,), \tag{9.69}$$
$$M^\dagger M = \mathrm{diag}(\, V_1^\dagger V_1,\ V_2^\dagger V_2,\ \cdots,\ V_n^\dagger V_n\,), \tag{9.70}$$

written in explicit forms, (9.69) *and* (9.70) *are equivalent to*

$$A_{i1}G_{1j} + A_{i2}G_{2j} + \cdots + A_{in}G_{nj} = \begin{cases} W_iW_i^\dagger & i = j \\ 0 & i \neq j \end{cases} \qquad i,\ j = 1,\ 2,\ \cdots,\ m,$$

$$G_{i1}A_{1j} + G_{i2}A_{2j} + \cdots + G_{im}A_{mj} = \begin{cases} V_i^\dagger V_i & i = j \\ 0 & i \neq j \end{cases} \qquad i,\ j = 1,\ 2,\ \cdots,\ n.$$

In addition to the expression given in (9.67) for $M^\dagger$, we can also derive some other expressions for G_{ij} in $M^\dagger$ from (9.32). But they are quite complicated in form, so we omit them here.

Various consequences can be derived from (9.67) when the submatrices in M satisfies some additional conditions, or M has some particular patterns, such as triangular forms, circulant forms and tridiagonal forms. Here we only give one consequences.

Corollary 9.19. *If the block matrix M in* (9.53) *satisfies the following rank additivity condition*

$$r(M) = \sum_{i=1}^{m}\sum_{j=1}^{n} r(A_{ij}), \tag{9.71}$$

then the Moore-Penrose inverse of M can be expressed as

$$M^\dagger = \begin{bmatrix} (E_{B_{11}}A_{11}F_{C_{11}})^\dagger & \cdots & (E_{B_{m1}}A_{m1}F_{C_{m1}})^\dagger \\ \vdots & & \vdots \\ (E_{B_{1n}}A_{1n}F_{C_{1n}})^\dagger & \cdots & (E_{B_{mn}}A_{mn}F_{C_{mn}})^\dagger \end{bmatrix}, \tag{9.72}$$

where B_{ij} and C_{ij} are defined in (9.56) *and* (9.57).

Proof. In fact, (9.71) is equivalent to

$$R(A_{ij}) \cap R(B_{ij}) = \{0\}, \qquad R(A_{ij}^*) \cap R(C_{ij}^*) = \{0\}, \qquad 1 \leq i \leq m,\ 1 \leq j \leq n,$$
$$R(C_{ij}) \cap R(D_{ij}) = \{0\}, \qquad R(B_{ij}^*) \cap R(D_{ij}^*) = \{0\}, \qquad 1 \leq i \leq m,\ 1 \leq j \leq n.$$

We can get from them $J(A_{ij}) = E_{B_{ij}}A_{ij}F_{C_{ij}}$. Putting them in (9.67) yields (9.72). □

The results so far established in the chapter are mainly based on the factorizations (9.1) and (9.12). However, if a given block matrix has certain special pattern such that we can factor it in some particular methods, then we can establish some new rank equalities through the special factorizations of the block matrix. Here we present some examples on the Moore-Penrose inverse of some special block matrices.

Theorem 9.20. *Let $A,\ B \in \mathcal{C}^{m\times n}$, $0 \neq p \in \mathcal{C}$, and let*

$$M = \begin{bmatrix} A & p^2B \\ B & A \end{bmatrix}, \qquad N = \begin{bmatrix} A + pB & 0 \\ 0 & A - pB \end{bmatrix}.$$

(a) *If $|p| = 1$, then*

$$\begin{bmatrix} A & p^2B \\ B & A \end{bmatrix}^\dagger = P_{2n}\begin{bmatrix} (\, A + pB\,)^\dagger & 0 \\ 0 & (\, A - pB\,)^\dagger \end{bmatrix} P_{2m}, \tag{9.73}$$

where

$$P_{2t} = P_{2t}^{-1} = \frac{1}{\sqrt{2}}\begin{bmatrix} I_t & pI_t \\ p^{-1}I_t & -I_t \end{bmatrix}, \qquad t = m,\ n.$$

In particular,

$$\begin{bmatrix} A & -B \\ B & A \end{bmatrix}^\dagger = \frac{1}{2}\begin{bmatrix} I_n & iI_n \\ -iI_n & -I_n \end{bmatrix}\begin{bmatrix} (\, A + iB\,)^\dagger & 0 \\ 0 & (\, A - iB\,)^\dagger \end{bmatrix}\begin{bmatrix} I_m & iI_m \\ -iI_m & -I_m \end{bmatrix}.$$

(b) *If $|p| \neq 1$, then*

$$r(\, M^\dagger - P_{2n}N^\dagger P_{2m}\,) = 2r\begin{bmatrix} A \\ B \end{bmatrix} + 2r[\, A,\ B\,] - 2r(\, A + pB\,) - 2r(\, A - pB\,). \tag{9.74}$$

(c) *Under* $|p| \neq 1$, (9.73) *holds if and only if* $R(A) \subseteq R(\,A + pB\,)$ *and* $R(A^*) \subseteq R(\,A^* - \overline{p}B^*\,)$.

Proof. It is easy to verify that the block matrix M can factor as $M = P_{2m}NP_{2n}$, where P_{2m} and P_{2n} are nonsingular with $P_{2m}^2 = I_{2m}$ and $P_{2n}^2 = I_{2n}$. In that case, we find by Theorem 8.14 that

$$r(\,M^\dagger - P_{2n}^{-1}N^\dagger P_{2m}^{-1}\,) = r(\,M^\dagger - P_{2n}N^\dagger P_{2m}\,) = r\begin{bmatrix} N \\ NP_{2n}P_{2n}^* \end{bmatrix} + r[\,N,\ P_{2m}^*P_{2m}N\,] - 2r(N), \quad (9.75)$$

where

$$P_{2m}^*P_{2m} = \frac{1}{2}\begin{bmatrix} I_m & \overline{p}^{-1}I_m \\ \overline{p}I_m & -I_m \end{bmatrix}\begin{bmatrix} I_m & pI_m \\ p^{-1}I_m & -I_m \end{bmatrix} = \frac{1}{2}\begin{bmatrix} (\,1+|p|^{-2}\,)I_m & (\,p - \overline{p}^{-1}\,)I_n \\ (\,\overline{p} - p^{-1}\,)I_m & (\,1+|p|^2\,)I_m \end{bmatrix},$$

$$P_{2n}P_{2n}^* = \frac{1}{2}\begin{bmatrix} I_n & pI_n \\ p^{-1}I_n & -I_n \end{bmatrix}\begin{bmatrix} I_n & \overline{p}^{-1}I_n \\ \overline{p}I_n & -I_n \end{bmatrix} = \frac{1}{2}\begin{bmatrix} (\,1+|p|^2\,)I_n & (\,\overline{p}^{-1} - p\,)I_n \\ (\,p^{-1} - \overline{p}\,)I_n & (\,1+|p|^{-2}\,)I_n \end{bmatrix}.$$

If $|p| = 1$, then $P_{2m}^*P_{2m} = I_{2m}$ and $P_{2n}P_{2n}^* = I_{2n}$. Thus the right-hand side of (9.75) becomes zero, which implies that $M^\dagger = P_{2n}N^\dagger P_{2m}$, the desired result in (9.73). If $|p| \neq 1$, then we find

$$r[\,N,\ P_{2m}^*P_{2m}N\,] = r\begin{bmatrix} 0 & A - pB & A + pB & 0 \\ A + pB & 0 & 0 & A - pB \end{bmatrix} = 2r[\,A,\ B\,].$$

Similarly $r\begin{bmatrix} N \\ NP_{2n}P_{2n}^* \end{bmatrix} = 2r\begin{bmatrix} A \\ B \end{bmatrix}$. Thus (9.75) becomes (9.74). □

Theorem 9.21. *Let* $M = \begin{bmatrix} A & A \\ A & A + B \end{bmatrix}$. *where* $A,\ B \in \mathcal{C}^{m\times n}$. *Then* M *factor as*

$$M = PNQ = \begin{bmatrix} I_m & 0 \\ I_m & I_m \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} I_n & I_n \\ 0 & I_n \end{bmatrix}.$$

In that case

$$r(\,M^\dagger - Q^{-1}N^\dagger P^{-1}\,) = 2r\begin{bmatrix} A \\ B \end{bmatrix} + 2r[\,A,\ B\,] - 2r(A) - 2r(B).$$

In particular, the Moore-Penrose inverse inverse of $M^\dagger$ *can be expressed as*

$$M^\dagger = Q^{-1}N^\dagger P^{-1} = \begin{bmatrix} A^\dagger + B^\dagger & -B^\dagger \\ -B^\dagger & B^\dagger \end{bmatrix},$$

if and only if $R(A) = R(B)$ *and* $R(A^*) = R(B^*)$.

We leave the verification of Theorem 9.21 to the reader. For the block matrix M in Theorem 9.21, we can also factor it, according to (9.12), as

$$M = PNQ = \begin{bmatrix} I_m & 0 \\ AA^\dagger & I_m \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} I_n & A^\dagger A \\ 0 & I_n \end{bmatrix}.$$

In that case

$$r(\,M^\dagger - Q^{-1}N^\dagger P^{-1}\,) = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\,A,\ B\,] - 2r(B).$$

Thus, we see that the Moore-Penrose inverse inverse of $M^\dagger$ can be expressed as

$$M^\dagger = Q^{-1}N^\dagger P^{-1} = \begin{bmatrix} A^\dagger + A^\dagger AB^\dagger AA^\dagger & -A^\dagger AB^\dagger \\ -B^\dagger AA^\dagger & B^\dagger \end{bmatrix},$$

if and only if $R(A) \subseteq R(B)$ and $R(A^*) \subseteq R(B^*)$.

Another interesting example is concerning the Moore-Penrose inverse of the $k \times k$ block matrix

$$M = \begin{bmatrix} A & B & \cdots & B \\ B & A & \cdots & B \\ \vdots & \vdots & \ddots & \vdots \\ B & B & \cdots & A \end{bmatrix}_{k\times k}, \quad (9.76)$$

where both A and B are $m \times n$ matrices. It is easy to verify that

$$M = P_m N Q_n = \begin{bmatrix} I_m & -I_m & \cdots & -I_m \\ I_m & I_m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ I_m & 0 & \cdots & I_m \end{bmatrix} \begin{bmatrix} A+(k-1)B & & & \\ & A-B & & \\ & & \ddots & \\ & & & A-B \end{bmatrix}$$
$$\times \begin{bmatrix} I_n/k & I_n/k & \cdots & I_n/k \\ -I_n/k & (k-1)I_n/k & \cdots & -I_n/k \\ \vdots & \vdots & \ddots & \vdots \\ -I_n/k & -I_n/k & \cdots & (k-1)I_n/k \end{bmatrix}, \tag{9.77}$$

where P and Q are nonsingular, and both of them satisfy

$$P^*P = \begin{bmatrix} kI_m & 0 & \cdots & 0 \\ 0 & 2I_m & \cdots & I_m \\ \vdots & \vdots & \ddots & \vdots \\ 0 & I_m & \cdots & 2I_m \end{bmatrix}, \quad QQ^* = \begin{bmatrix} I_n/k & 0 & \cdots & 0 \\ 0 & (k-1)I_n/k & \cdots & -I_n/k \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -I_n/k & \cdots & (k-1)I_n/k \end{bmatrix}.$$

Note that N is diagonal. Then it is easy to verify that $r[\, N, \; P^*PN \,] = r(N)$ and $r\begin{bmatrix} N \\ NQQ^* \end{bmatrix} = r(N)$. Thus we have $M^\dagger = Q_n^{-1}N^\dagger P_m^{-1}$ according to (8.20). Furthermore, one can verify that $P_mQ_m = I_{km}$ and $P_nQ_n = I_{kn}$ when $m = n$. Hence we can write $M^\dagger$ as $M^\dagger = P_nN^\dagger Q_m$. Written in an explicit form

$$M^\dagger = \begin{bmatrix} S & T & \cdots & T \\ T & S & \cdots & T \\ \vdots & \vdots & \ddots & \vdots \\ T & T & \cdots & S \end{bmatrix}_{k\times k}, \tag{9.78}$$

where

$$S = \frac{1}{k}[\, A+(k-1)B \,]^\dagger + \frac{k-1}{k}(\, A-B \,)^\dagger, \qquad T = \frac{1}{k}[\, A+(k-1)B \,]^\dagger - \frac{1}{k}(\, A-B \,)^\dagger. \tag{9.79}$$

The expression (9.78) illustrates that $M^\dagger$ has the same pattern as M. To find $M^\dagger$, what we actually need to do is to find $[\, A+(k-1)B \,]^\dagger$ and $(\, A-B \,)^\dagger$, and then put them in (9.78). Some interesting subsequent results can be derived from (9.77) and (9.78). For example

$$\begin{aligned} r(\, MM^\dagger - M^\dagger M \,) = {} & 2r[\, A+(k-1)B, \; (\, A+(k-1)B \,)^* \,] + 2(k-1)r[\, A-B, \; (\, A-B \,)^* \,] - 2r[\, A+(k-1)B \,] \\ & -2(k-1)r(\, A-B \,). \end{aligned} \tag{9.80}$$

In particular, M in (9.76) is EP if and only if both $A+(k-1)B$ and $A-B$ are EP. We leave the verification of the result to the reader. One can also find $r[\, (MM^\dagger)(M^\dagger M) - (M^\dagger M)(MM^\dagger) \,]$ and $r(\, M^*M^\dagger - M^\dagger M^* \,)$ and so on for M in (9.76).

A more general work than (9.78) is to consider Moore-Penrose inverses of block circulant matrices. This topic was examined by Smith in [126] and some nice properties on Moore-Penrose inverses of block circulant matrices were presented there. Much similar to what we have done for M in (9.76), through block factorization, we can also simply find a general expression for Moore-Penrose inverses of block circulant matrices, and derive from them various consequences. We shall present the corresponding results in Chapter 11.

Remark. It should be pointed out that many results similar to those in Theorems 9.20 and 9.21, as well as in (9.78), (9.83)—(9.86), can be trvially established. In fact, properly choosing block matrices P, N and Q and then applying Theorems 8.12, 8.13 and 8.14 to them, one can find out various rank equalities related to the Moore-Penrose inverses of the block matrices. Based on those rank equalities, one can further derive necessary and sufficient conditions for $(PNQ)^\dagger = Q^\dagger N^\dagger P^\dagger$ or $(PNQ)^\dagger = Q^{-1}N^\dagger P^{-1}$ to

hold for these block matrices. We hope the reader to try this method and find some more interesting or unexpected conclusions about Moore-Penrose inverses of block matrices.

In addition to the methods mentioned above for finding Moore-Penrose inverses of block matrices, another possible tool is the identity (8.25) for the Moore-Penrose inverse of product of three matrices.

Chapter 10

Rank equalities for Moore-Penrose inverses of sums of matrices

In this chapter, we establish rank equalities related to Moore-Penrose inverses of sums of matrices and consider their various consequences.

Theorem 10.1. *Let $A, B \in \mathcal{C}^{m\times n}$ be given and let $N = A + B$. Then*

$$r[\, N - N(\, A^{\dagger} + B^{\dagger}\,)N\,] = r\begin{bmatrix} AB^*A & AA^*B + AB^*B \\ BA^*A + BB^*A & BA^*B \end{bmatrix} + r(N) - r(A) - r(B). \tag{10.1}$$

In particular,

$$A^{\dagger} + B^{\dagger} \in \{(A+B)^-\} \Leftrightarrow r\begin{bmatrix} AB^*A & AA^*B + AB^*B \\ BA^*A + BB^*A & BA^*B \end{bmatrix} = r(A) + r(B) - r(N). \tag{10.2}$$

Proof. It follows by (2.2) and block elementary operations that

$$\begin{aligned} & r[\, N - N(\, A^{\dagger} + B^{\dagger}\,)N\,] \\ &= r\begin{bmatrix} A^*AA^* & 0 & A^*N \\ 0 & B^*BB^* & B^*N \\ NA^* & NB^* & N \end{bmatrix} - r(A) - r(B) \\ &= r\begin{bmatrix} -A^*BA^* & -A^*AB^* - A^*BB^* & 0 \\ -B^*AA^* - B^*BA^* & -B^*AB^* & 0 \\ 0 & 0 & N \end{bmatrix} - r(A) - r(B) \\ &= r\begin{bmatrix} AB^*A & AA^*B + AB^*B \\ BA^*A + BB^*A & BA^*B \end{bmatrix} + r(N) - r(A) - r(B). \end{aligned}$$

Thus we have (10.1) and (10.2). $\square$

A general result is given below.

Theorem 10.2. *Let $A_1, \ , A_2, \cdots, A_k \in \mathcal{C}^{m\times n}$ be given and let $A = A_1 + A_2 + \cdots + A_k$, $X = A_1^{\dagger} + A_2^{\dagger} + \cdots + A_k^{\dagger}$. Then*

$$r(\, A - AXA\,) = r(\, DD^*D - PA^*Q\,) - r(D) + r(A), \tag{10.3}$$

where

$$D = \mathrm{diag}(\, A_1, \, A_2, \cdots, \, A_k\,), \quad P^* = [\, A_1^*, \ A_2^*, \ \cdots, \ A_k^*\,], \quad Q = [\, A_1, \, A_2, \cdots, \, A_k\,].$$

In particular,

$$X \in \{A^-\} \Leftrightarrow r(\, DD^*D - PA^*Q\,) = r(D) - r(A), \quad i.e., \quad PA^*Q \leq_{rs} DD^*D. \tag{10.4}$$

Proof. Let $P_1 = [\, I_n, \cdots, I_n\,]$ and $Q_1 = [\, I_m, \cdots, I_m\,]^*$. Then $X = P_1D^{\dagger}Q_1$. In that case, it follows by (2.1) that

$$r(\, A - AXA\,) \ = \ r(\, A - AP_1D^{\dagger}Q_1A\,)$$

$$
\begin{aligned}
&= r\begin{bmatrix} D^*DD^* & D^*Q_1A \\ AP_1D^* & A \end{bmatrix} - r(D) \\
&= r\begin{bmatrix} D^*DD^* - D^*Q_1AP_1D^* & 0 \\ 0 & A \end{bmatrix} - r(D) \\
&= r(\, D^*DD^* - D^*Q_1AP_1D^* \,) + r(A) - r(D) \\
&= r(\, DD^*D - DP_1^*A^*Q_1^*D \,) + r(A) - r(D)
\end{aligned}
$$

as required for (10.3). □

Theorem 10.3. *Let $A, B \in \mathcal{C}^{m\times n}$ be given and let $N = A + B$. Then*

$$
r(\, N^\dagger - A^\dagger - B^\dagger \,) = r\begin{bmatrix} -NN^*N & 0 & 0 & N \\ 0 & AA^*A & 0 & A \\ 0 & 0 & BB^*B & B \\ N & A & B & 0 \end{bmatrix} - r(N) - r(A) - r(B). \tag{10.5}
$$

In particular,

$$
N^\dagger = A^\dagger + B^\dagger \Leftrightarrow r\begin{bmatrix} -NN^*N & 0 & 0 & N \\ 0 & AA^*A & 0 & A \\ 0 & 0 & BB^*B & B \\ N & A & B & 0 \end{bmatrix} = r(N) + r(A) + r(B). \tag{10.6}
$$

Proof. Follows immediately from (2.7). □

It is well known that for any two nonsingular matrices A and B, there always is $A(\, A^{-1} + B^{-1} \,)B = A + B$. Now for Moore-Penrose inverses of matrices we have the following.

Theorem 10.4. *Let $A, B \in \mathcal{C}^{m\times n}$ be given. Then*

$$
r[\, A + B - A(\, A^\dagger + B^\dagger \,)B \,] = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A, \; B \,] - r(A) - r(B), \tag{10.7}
$$

and

$$
r[\, A^\dagger + B^\dagger - A^\dagger(A + B)B^\dagger \,] = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A, \; B \,] - r(A) - r(B). \tag{10.8}
$$

In particular,

$$
A(\, A^\dagger + B^\dagger \,)B = A + B \Leftrightarrow A^\dagger(A + B)B^\dagger = A^\dagger + B^\dagger \Leftrightarrow R(A) = R(B) \;\; \textit{and} \;\; R(A^*) = R(B^*). \tag{10.9}
$$

Proof. Writing

$$
A + B - A(\, A^\dagger + B^\dagger \,)B = A + B - [\, A, \; A \,]\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^\dagger \begin{bmatrix} B \\ B \end{bmatrix}
$$

and then applying (2.1) to it produce (10.7). Replacing A and B in (10.7) respectively by $A^\dagger$ and $B^\dagger$ leads to (10.8). The equivalences in (10.9) follow from (10.7) and (10.8). □

Theorem 10.5. *Let $A, B \in \mathcal{C}^{m\times n}$ be given and let $N = A + B$. Then*

$$
r[\, N - N(\, (E_BAF_B)^\dagger + (E_ABF_A)^\dagger \,)N \,] = r(N) + 2r(A) + 2r(B) - r\begin{bmatrix} A & B \\ B & 0 \end{bmatrix} - r\begin{bmatrix} B & A \\ A & 0 \end{bmatrix}. \tag{10.10}
$$

In particular,

$$
(E_BAF_B)^\dagger + (E_ABF_A)^\dagger \in \{(A + B)^-\} \Leftrightarrow r(\, A + B \,) = r(E_BAF_B) + r(E_ABF_A). \tag{10.11}
$$

Proof. Let $P = E_BAF_B$ and $Q = E_ABF_A$. Then it is easy to verify that

$$
P^*B = 0, \quad BP^* = 0, \quad Q^*A = 0, \quad AQ^* = 0, \quad P^*PP^* = P^*AP^*, \quad Q^*QQ^* = Q^*BQ^*.
$$

Thus we find by (2.2) that

$$\begin{aligned} r[\,N - N(\,P^{\dagger} + Q^{\dagger}\,)N\,] &= r\begin{bmatrix} P^*PP^* & 0 & P^*N \\ 0 & Q^*QQ^* & Q^*N \\ NP^* & NQ^* & N \end{bmatrix} - r(P) - r(Q) \\ &= r\begin{bmatrix} P^*AP^* & 0 & P^*A \\ 0 & Q^*BQ^* & Q^*B \\ AP^* & BQ^* & A+B \end{bmatrix} - r(P) - r(Q) \\ &= r\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & A+B \end{bmatrix} - r(P) - r(Q) = r(N) - r(P) - r(Q), \end{aligned}$$

where

$$r(P) = r\begin{bmatrix} A & B \\ B & 0 \end{bmatrix} - 2r(B), \qquad r(Q) = r\begin{bmatrix} B & A \\ A & 0 \end{bmatrix} - 2r(A),$$

by (1.4). Hence we have (10.10) and (10.11). □

Theorem 10.6. *Let $A,\ B \in \mathcal{C}^{m\times n}$ be given. Then*

(a) $r(\,(\,A+B\,)(\,A+B\,)^{\dagger} - [\,A,\ B\,][\,A,\ B\,]^{\dagger}\,) = r[\,A,\ B\,] - r(\,A+B\,)$.

(b) $r\left(\,(\,A+B\,)^{\dagger}(\,A+B\,) - \begin{bmatrix} A \\ B \end{bmatrix}^{\dagger}\begin{bmatrix} A \\ B \end{bmatrix}\right) = r\begin{bmatrix} A \\ B \end{bmatrix} - r(\,A+B\,)$.

In particular,

(c) $(\,A+B\,)(\,A+B\,)^{\dagger} = [\,A,\ B\,][\,A,\ B\,]^{\dagger} \Leftrightarrow r[\,A,\ B\,] = r(\,A+B\,) \Leftrightarrow R(A) \subseteq R(\,A+B\,)$ *and* $R(B) \subseteq R(\,A+B\,)$.

(d) $(\,A+B\,)^{\dagger}(\,A+B\,) = \begin{bmatrix} A \\ B \end{bmatrix}^{\dagger}\begin{bmatrix} A \\ B \end{bmatrix} \Leftrightarrow r\begin{bmatrix} A \\ B \end{bmatrix} = r(\,A+B\,) \Leftrightarrow R(A^*) \subseteq R(\,A^*+B^*\,)$ *and* $R(B^*) \subseteq R(\,A^*+B^*\,)$.

Proof. Let $N = A+B$ and $M = [\,A,\ B\,]$. Then it follows from Theorem 7.2(a) that

$$\begin{aligned} r(\,NN^{\dagger} - MM^{\dagger}\,) &= 2r[\,N,\ M\,] - r(N) - r(M) \\ &= 2r[\,A+B,\ A,\ B\,] - r(\,A+B\,) - r(A) - r(B) \\ &= r[\,A,\ B\,] - r(A) - r(B), \end{aligned}$$

as required for Part (a). Similarly we have Part (b). □

In general we have the following.

Theorem 10.7. *Let $A_1,\ A_2,\ \cdots,\ A_k \in \mathcal{C}^{m\times n}$ be given and let $A = A_1+A_2+\cdots+A_k$, $M = [\,A_1,\ A_2,\ \cdots,\ A_k\,]$ and $N^* = [\,A_1^*,\ A_2^*,\ \cdots,\ A_k^*\,]$. Then*

(a) $r(\,AA^{\dagger} - MM^{\dagger}\,) = r(M) - r(A)$.

(b) $r(\,A^{\dagger}A - N^{\dagger}N\,) = r(N) - r(A)$.

(c) $AA^{\dagger} = MM^{\dagger} \Leftrightarrow r(M) = r(A) \Leftrightarrow R(A_i) \subseteq R(M),\ i = 1,\ 2,\ \cdots,\ k$.

(d) $A^{\dagger}A = N^{\dagger}N \Leftrightarrow r(N) = r(A) \Leftrightarrow R(A_i^*) \subseteq R(N^*),\ i = 1,\ 2,\ \cdots,\ k$.

Theorem 10.8. *Let $A,\ B \in \mathcal{C}^{m\times n}$ be given and let $N = A+B$. Then*

(a) $r(AN^{\dagger}B) = r(NA^*) + r(B^*N) - r(N)$.

(b) $r(\,AN^{\dagger}B\,) = r(A) + r(B) - r(N)$, *if* $R(A^*) \subseteq R(N^*)$ *and* $R(B) \subseteq R(N)$.

(c) $r(\,AN^{\dagger}B - BN^{\dagger}A\,) = r\begin{bmatrix} N \\ NA^* \end{bmatrix} + r[\,N,\ AN^*\,] - 2r(N)$.

(d) $AN^{\dagger}B = 0 \Leftrightarrow r(NA^*) + r(B^*N) = r(N)$.

(e) $AN^{\dagger}B = BN^{\dagger}A \Leftrightarrow R(AN^*) \subseteq R(N)$ *and* $R(A^*N) \subseteq R(N^*)$.

(f) $AN^{\dagger}B = BN^{\dagger}A$, *if* $R(A) \subseteq R(N)$ *and* $R(A^*) \subseteq R(N^*)$.

Proof. It follows by (2.1) that

$$r(AN^{\dagger}B) \;=\; r\begin{bmatrix} N^*NN^* & N^*B \\ AN^* & 0 \end{bmatrix} - r(N)$$

$$
\begin{aligned}
&= r\begin{bmatrix} N^*AN^* + N^*BN^* & N^*B \\ AN^* & 0 \end{bmatrix} - r(N) \\
&= r\begin{bmatrix} 0 & N^*B \\ AN^* & 0 \end{bmatrix} - r(N) = r(NA^*) + r(B^*N) - r(N),
\end{aligned}
$$

as required for Part (a). Under $R(A^*) \subseteq R(N^*)$ and $R(B) \subseteq R(N)$, we know that $r(NA^*) = r(A)$ and $r(B^*N) = r(B)$. Thus we have Part (b). Similarly it follows by (2.1) that

$$
\begin{aligned}
&r(\, AN^\dagger B - BN^\dagger A\,) \\
&= r\left([\,A,\ B\,] \begin{bmatrix} N & 0 \\ 0 & -N \end{bmatrix}^\dagger \begin{bmatrix} B \\ A \end{bmatrix} \right) \\
&= r\begin{bmatrix} N^*NN^* & 0 & N^*B \\ 0 & -N^*NN^* & N^*A \\ AN^* & BN^* & 0 \end{bmatrix} - 2r(N) \\
&= r\begin{bmatrix} N^*AN^* & N^*BN^* & N^*B \\ -N^*AN^* & -N^*BN^* & N^*A \\ AN^* & BN^* & 0 \end{bmatrix} - 2r(N) \\
&= r\begin{bmatrix} 0 & 0 & N^*B \\ 0 & 0 & N^*A \\ AN^* & BN^* & 0 \end{bmatrix} - 2r(N) \\
&= r\begin{bmatrix} N^*B \\ N^*A \end{bmatrix} + r[\,AN^*,\ BN^*\,] - 2r(N) \\
&= r\begin{bmatrix} N^*N \\ N^*A \end{bmatrix} + r[\,AN^*,\ NN^*\,] - 2r(N) = r\begin{bmatrix} N \\ N^*A \end{bmatrix} + r[\,AN^*,\ N\,] - 2r(N),
\end{aligned}
$$

as required for Part (c). □

It is well known that if $R(A^*) \subseteq R(N^*)$ and $R(B) \subseteq R(N)$, the product $A(\,A + B\,)^\dagger B$ is called the parallel sum of A and B and often denoted by $P(A,\, B)$. The results in Theorem 10.8(b) and (f) show that if A and B are parallel summable, then

$$
r[\,P(A,\, B)\,] = r(A) + r(B) - r(\,A + B\,) \quad \text{and} \quad P(A,\, B) = P(B,\, A).
$$

These two properties were obtained by Rao and Mitra [118] with a different method.

The following three theorems are derived directly from (2.1). Their proofs are omitted here.

Theorem 10.9. *Let $A,\, B \in \mathcal{C}^{m\times n}$ be given. Then*

$$
r\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} - \begin{bmatrix} A \\ B \end{bmatrix} (\,A + B\,)^\dagger [\,A,\ B\,] \right) = r(A) + r(B) - r(\,A + B\,). \tag{10.12}
$$

In particular,

$$
\begin{bmatrix} A \\ B \end{bmatrix} (\,A + B\,)^\dagger [\,A,\ B\,] = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Leftrightarrow r(\,A + B\,) = r(A) + r(B). \tag{10.13}
$$

The equivalence in (10.13) was established by Marsaglia and Styan [83].

Theorem 10.10. *Let $A_1,\, A_2,\, \cdots,\, A_k \in \mathcal{C}^{m\times n}$ be given and denote*

$$
A = \mathrm{diag}(\,A_1,\ A_2,\ \cdots,\ A_k\,), \qquad N = A_1 + A_2 + \cdots + A_k.
$$

Then

$$
r\left(A - \begin{bmatrix} A_1 \\ \vdots \\ A_k \end{bmatrix} N^\dagger [\,A_1,\, \cdots,\, A_k\,] \right) = r(A_1) + \cdots + r(A_k) - r(N). \tag{10.14}
$$

In particular,

$$\begin{bmatrix} A_1 \\ \vdots \\ A_k \end{bmatrix} N^\dagger [\, A_1, \, \cdots, \, A_k \,] = A \Leftrightarrow r(N) = r(A_1) + \cdots + r(A_k). \tag{10.15}$$

The equivalence in (10.15) was established by Marsaglia and Styan [83].

Theorem 10.11. *Let $A, \ B \in \mathcal{C}^{m \times n}$ be given and let $N = A + B$. Then*

(a) $r(\, A - AN^\dagger A) = r(NB^*N) + r(A) - r(N)$.

(b) $r(\, A - AN^\dagger A\,) = r(A) + r(B) - r(N)$, *if* $R(A) \subseteq R(N)$ *and* $R(A^*) \subseteq R(N^*)$.

(c) $N^\dagger \in \{A^-\} \ \Leftrightarrow \ r(NB^*N) = r(N) - r(A)$.

(d) $N^\dagger \in \{A^-\}$ *if* $r(N) = r(A) + r(B)$.

Proof. Immediate by (2.1). □

Notice that the sum $A + B$ can be expressed as

$$A + B = [\, A, \ I\,] \begin{bmatrix} I \\ B \end{bmatrix} = [\, AA^\dagger, \ B\,] \begin{bmatrix} A \\ B^\dagger B \end{bmatrix} = [\, A, \ B\,] \begin{bmatrix} A^\dagger A \\ B^\dagger B \end{bmatrix} = [\, A, \ B\,] \begin{bmatrix} A^\dagger & 0 \\ 0 & B^\dagger \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}.$$

Through the results in Chapter 8, we can also establish some rank equalities for sums of matrices.

Theorem 10.12. *Let $A, \ B \in \mathcal{C}^{m \times n}$ be given. Then*

(a) $r\left((\, A + B\,) - (\, A + B\,) \begin{bmatrix} I_n \\ B \end{bmatrix}^\dagger [\, A, \ I_m\,]^\dagger (\, A + B\,) \right) = r(\, I_n - A^*B\,) + r(\, A + B\,) - n.$

(b) $\begin{bmatrix} I_n \\ B \end{bmatrix}^\dagger [\, A, \ I_m\,]^\dagger \subseteq \{\, (\, A + B\,)^- \,\} \Leftrightarrow r(\, I_n - A^*B\,) + r(\, A + B\,) = n.$

Proof. Write $A + B = [\, A, \ I_m\,] \begin{bmatrix} I_n \\ B \end{bmatrix} = PQ$. Then we find by (8.1) that

$$\begin{aligned} r(\, PQ - PQQ^\dagger P^\dagger PQ\,) &= r[\, P^*, \ Q\,] + r(PQ) - r(P) - r(Q) \\ &= r\begin{bmatrix} A^* & I_n \\ I_m & B \end{bmatrix} + r(\, A + B\,) - r[\, A, \ I_m\,] - r\begin{bmatrix} I_n \\ B \end{bmatrix} \\ &= r(\, I_n - A^*B\,) + r(\, A + B\,) - n. \end{aligned}$$

Parts (a) and (b) follow from it. □

Theorem 10.13. *Let $A, \ B \in \mathcal{C}^{m \times n}$ be given. Then*

(a) $r\left((\, A + B\,) - (\, A + B\,) \begin{bmatrix} A \\ B^\dagger B \end{bmatrix}^\dagger [\, AA^\dagger, \ B\,]^\dagger (\, A + B\,) \right) = r\begin{bmatrix} A^* & A^*A \\ BB^* & B \end{bmatrix} + r(\, A + B\,) - r[\, A, \ B\,] - r\begin{bmatrix} A \\ B \end{bmatrix}.$

(b) $r\left((\, A + B\,) - (\, A + B\,) \begin{bmatrix} A \\ B \end{bmatrix}^\dagger [\, AA^\dagger, \ BB^\dagger\,]^\dagger (\, A + B\,) \right) = r\begin{bmatrix} A & B \\ A^*A & B^*B \end{bmatrix} + r(\, A + B\,) - r[\, A, \ B\,] - r\begin{bmatrix} A \\ B \end{bmatrix}.$

(c) $\begin{bmatrix} A \\ B^\dagger B \end{bmatrix}^\dagger [\, AA^\dagger, \ B\,]^\dagger \subseteq \{\, (\, A + B\,)^- \,\} \Leftrightarrow r\begin{bmatrix} A^* & A^*A \\ BB^* & B \end{bmatrix} = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A, \ B\,] - r(\, A + B\,).$

(d) $\begin{bmatrix} A \\ B \end{bmatrix}^\dagger [\, AA^\dagger, \ BB^\dagger\,]^\dagger \subseteq \{\, (\, A + B\,)^- \,\} \Leftrightarrow r\begin{bmatrix} A & B \\ A^*A & B^*B \end{bmatrix} = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A, \ B\,] - r(\, A + B\,).$

Proof. Writing $A + B = [\, AA^\dagger, \ B\,] \begin{bmatrix} A \\ B^\dagger B \end{bmatrix} = [\, AA^\dagger, \ BB^\dagger\,] \begin{bmatrix} A \\ B \end{bmatrix}$, and then applying (8.1) to them yields Parts (a) and (b). □

Theorem 10.14. *Let $A,\ B \in \mathcal{C}^{m\times n}$ be given. Then*

$$(A + B)^{\dagger} = \begin{bmatrix} I_n \\ B \end{bmatrix}^{\dagger} [\, A,\ I_m\,]^{\dagger} \tag{10.15}$$

holds if and only if

$$(I_m - BA^*)(A + B) = (A + B)(I_n - B^*A) = 0. \tag{10.16}$$

Proof. Write $A + B = [\, A,\ I_m\,] \begin{bmatrix} I_n \\ B \end{bmatrix} = PQ$. Then we find by Theorem 8.2(e) that $(PQ)^{\dagger} = Q^{\dagger}P^{\dagger}$ if and only if

$$r[\, P^*PQ,\ Q\,] = r(Q), \quad \text{and} \quad r\begin{bmatrix} P \\ PQQ^* \end{bmatrix} = r(P). \tag{10.17}$$

Notice that $r(Q) = n$ and $r(P) = m$, and

$$r[\, P^*PQ,\ Q\,] = r\begin{bmatrix} A^*(A + B) & I_n \\ A + B & B \end{bmatrix} = n + r[\, (I_m - BA^*)(A + B)\,],$$

$$r\begin{bmatrix} P \\ PQQ^* \end{bmatrix} = r\begin{bmatrix} A & I_m \\ A + B & (A + B)B^* \end{bmatrix} = m + r[\, (A + B)(I_n - B^*A)\,].$$

In that case, (10.17) reduces to (10.16). □

Theorem 10.15. *Let $A,\ B \in \mathcal{C}^{m\times n}$ be given. Then*

$$\begin{aligned} & r\left((A + B) - (A + B) \begin{bmatrix} A \\ B \end{bmatrix}^{\dagger} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} [\, A,\ B\,]^{\dagger} (A + B) \right) \\ & = r\begin{bmatrix} A + B & AA^* + BB^* \\ A^*A + B^*B & A^*AA^* + B^*BB^* \end{bmatrix} + r(A + B) - r[\, A,\ B\,] - r\begin{bmatrix} A \\ B \end{bmatrix}. \end{aligned} \tag{10.18}$$

In particular,

$$\begin{bmatrix} A \\ B \end{bmatrix}^{\dagger} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} [\, A,\ B\,]^{\dagger} \subseteq \{\, (A + B)^- \,\}$$

holds if and only if

$$r\begin{bmatrix} A + B & AA^* + BB^* \\ A^*A + B^*B & A^*AA^* + B^*BB^* \end{bmatrix} = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A,\ B\,] - r(A + B). \tag{10.19}$$

Proof. Writing $A + B = [\, I,\ I\,] \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} I \\ I \end{bmatrix} = PNQ$, and then applying (8.2) to it yields (10.18). □

Theorem 10.16. *Let $A,\ B \in \mathcal{C}^{m\times n}$ be given. Then*

$$\begin{aligned} & r\left((A + B)^{\dagger} - \begin{bmatrix} A \\ B \end{bmatrix}^{\dagger} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} [\, A,\ B\,]^{\dagger} \right) \\ & = r\left(\begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{bmatrix} A^* & 0 \\ 0 & B^* \end{bmatrix} \begin{bmatrix} A & B \\ B & A \end{bmatrix} \right) - r(A + B). \end{aligned} \tag{10.20}$$

In particular,

$$(A + B)^{\dagger} = \begin{bmatrix} A \\ B \end{bmatrix}^{\dagger} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} [\, A,\ B\,]^{\dagger} \Leftrightarrow r\left(\begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{bmatrix} A^* & 0 \\ 0 & B^* \end{bmatrix} \begin{bmatrix} A & B \\ B & A \end{bmatrix} \right) = r(A + B). \tag{10.21}$$

Proof. Writing $A + B = [\, I,\ I\,] \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} I \\ I \end{bmatrix} = PNQ$, and then applying (8.4) to it yields (10.20). □

The above several results can also be extended to sums of k matrices. In the remainder of this chapter, we present a set of results related to expressions of Moore-Penrose inverses of Schur complements. These results have appeared in the author's recent paper [136].

Theorem 10.17. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$, $C \in \mathcal{C}^{l\times n}$ and $D \in \mathcal{C}^{l\times n}$ be given, and satisfy the rank additivity condition*

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r\begin{bmatrix} B \\ D \end{bmatrix} = r[\,A,\ B\,] + r[\,C,\ D\,]. \tag{10.22}$$

Then the following inversion formula holds

$$\begin{aligned}(E_{B_2}S_DF_{C_2})^{\dagger} = {} & A^{\dagger} + A^{\dagger}BJ^{\dagger}(D)CA^{\dagger} + C_1^{\dagger}[\,S_AJ^{\dagger}(D)S_A - S_A\,]B_1^{\dagger} \\ & - A^{\dagger}B[\,I - J^{\dagger}(D)S_A\,]B_1^{\dagger} - C_1^{\dagger}[\,I - S_AJ^{\dagger}(D)\,]CA^{\dagger},\end{aligned} \tag{10.23}$$

where

$$S_A = D - CA^{\dagger}B, \qquad S_D = A - BD^{\dagger}C, \qquad J(D) = E_{C_1}S_AF_{B_1},$$
$$B_1 = E_AB, \qquad B_2 = BF_D, \qquad C_1 = CF_A, \qquad C_2 = E_DC.$$

Proof. Follows immediately from the two expressions of $M^{\dagger}$ in Theorem 9.8. $\square$

The results given below are all the special cases of the general formula (10.23).

Corollary 10.18. *If A, B, C and D satisfy*

$$R\begin{bmatrix} A \\ 0 \end{bmatrix} \subseteq R\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \qquad \text{and} \qquad R\begin{bmatrix} A^* \\ 0 \end{bmatrix} \subseteq R\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix}, \tag{10.24}$$

and the following two conditions

$$R(CS_D^*) \subseteq R(D), \qquad R(B^*S_D) \subseteq R(D^*), \tag{10.25}$$

or more specifically satisfy the four conditions

$$R(C) \subseteq R(D), \quad R(B^*) \subseteq R(D^*), \quad R(B) \subseteq R(S_D), \quad R(C^*) \subseteq R(S_D^*), \tag{10.26}$$

then the Moore-Penrose inverse of the Schur complement $S_D = A - BD^{\dagger}C$ satisfies the inversion formula

$$\begin{aligned}(A - BD^{\dagger}C)^{\dagger} = {} & A^{\dagger} + A^{\dagger}BJ^{\dagger}(D)CA^{\dagger} + C_1^{\dagger}[\,S_AJ^{\dagger}(D)S_A - S_A\,]B_1^{\dagger} \\ & - A^{\dagger}B[\,I - J^{\dagger}(D)S_A\,]B_1^{\dagger} - C_1^{\dagger}[\,I - S_AJ^{\dagger}(D)\,]CA^{\dagger},\end{aligned} \tag{10.27}$$

where S_A, B_1, C_1 and $J(D)$ are defined in (10.23).

Proof. It is obvious that (10.25) is equivalent to $(E_DC)S_D^* = 0$ and $S_D^*(BF_D) = 0$, or equivalently

$$S_D(E_DC)^{\dagger} = 0 \qquad \text{and} \qquad (BF_D)^{\dagger}S_D = 0. \tag{10.28}$$

These two equalities clearly imply that S_D, E_DC and BF_D satisfy (9.31). Hence by Lemma 9.7, we know that under (10.24) and (10.25), A, B, C and D naturally satisfy (10.22). Now substituting (10.28) into the left-hand side of (10.23) yields $J^{\dagger}(A) = (A - BD^{\dagger}C)^{\dagger}$. Hence (10.23) becomes (10.27). Observe that (10.28) is a special case of (10.25), hence (10.27) is also true under (10.26). $\square$

Corollary 10.19. *If A, B, C and D satisfy* (10.27), (10.25) *and the following two conditions*

$$R(CF_A) \cap R(S_A) = \{0\} \quad \text{and} \quad R[(E_AB)^*] \cap R(S_A^*) = \{0\}, \tag{10.29}$$

then

$$(A - BD^{\dagger}C)^{\dagger} = A^{\dagger} + A^{\dagger}BJ^{\dagger}(D)CA^{\dagger} - A^{\dagger}B[\,I - J^{\dagger}(D)S_A\,]B_1^{\dagger} - C_1^{\dagger}[\,I - S_AJ^{\dagger}(D)\,]CA^{\dagger}, \tag{10.30}$$

where S_A, B_1, C_1 and $J(D)$ are defined in (10.23).

Proof. According to Theorem 7.8, the two conditions in (10.27) imply that $S_A J^{\dagger}(D) S_A = S_A$. Hence (10.27) is simplified to (10.30). $\square$

Corollary 10.20. *If A, B, C and D satisfy* (10.24), (10.25) *and the following two conditions*

$$\mathscr{R}(BS_A^*) \subseteq \mathscr{R}(A) \quad and \quad \mathscr{R}(C^*S_A) \subseteq \mathscr{R}(A^*), \tag{10.31}$$

then

$$(A - BD^{\dagger}C)^{\dagger} = A^{\dagger} + A^{\dagger}BS_A^{\dagger}CA^{\dagger} - A^{\dagger}B(E_AB)^{\dagger} - (CF_A)^{\dagger}CA^{\dagger}. \tag{10.32}$$

where $S_A = D - CA^{\dagger}B$.

Proof. Clearly, (10.31) is equivalent to $(E_AB)S_A^* = 0$ and $S_A^*(CF_A) = 0$, which can also equivalently be expressed as $S_A(E_AB)^{\dagger} = 0$ and $(CF_A)^{\dagger}S_A = 0$. In that case, $J(D) = E_{C_1}S_AF_{B_1} = S_A$. Hence (10.27) is simplified to (10.32). $\square$

Corollary 10.21. *If A, B, C and D satisfy* (10.24), (10.25) *and the following two conditions*

$$\mathscr{R}(B) \subseteq \mathscr{R}(A) \quad and \quad \mathscr{R}(C^*) \subseteq \mathscr{R}(A^*), \tag{10.33}$$

then

$$(A - BD^{\dagger}C)^{\dagger} = A^{\dagger} + A^{\dagger}B(D - CA^{\dagger}B)^{\dagger}CA^{\dagger}. \tag{10.34}$$

Proof. The two inclusions in (10.33) are equivalent to $E_AB = 0$ and $CF_A = 0$. Substituting them into (10.27) yields (10.34). $\square$

Corollary 10.22. *If A, B, C and D satisfy the following four conditions*

$$\mathscr{R}(A) \cap \mathscr{R}(B) = \{0\}, \quad \mathscr{R}(A^*) \cap \mathscr{R}(C^*) = \{0\}, \quad \mathscr{R}(C) = \mathscr{R}(D), \quad \mathscr{R}(B^*) = \mathscr{R}(D^*), \tag{10.35}$$

then

$$(A - BD^{\dagger}C)^{\dagger} = A^{\dagger} - A^{\dagger}B(E_AB)^{\dagger} - (CF_A)^{\dagger}CA^{\dagger} + (CF_A)^{\dagger}S_A(E_AB)^{\dagger}. \tag{10.36}$$

Proof. Under (10.35), A, B, C and D naturally satisfy the rank additivity condition in (10.22). Besides, from (10.35) and Theorem 7.2(c) and (d) we can derive

$$B_1^{\dagger}B_1 = B^{\dagger}B, \quad C_1C_1^{\dagger} = CC^{\dagger}, \quad B_2 = 0, \quad C_2 = 0, \quad J(D) = 0.$$

Substituting them into (10.23) yields (10.36). $\square$

If D is invertible, or $D = I$, or $B = C = -D$, then the inversion formula (10.23) can reduce to some other simpler forms. For simplicity, we do not list them here.

Chapter 11

Moore-Penrose inverses of block circulant matrices

Inverses or Moore-Penrose inverses of circulant matrices and block circulant matrices is an attractive topic in matrix theory and lots of results can be find in the literature (see, e.g., [39, 40, 122, 126, 127, 141]). To find the general expression for inverses or Moore-Penrose inverses of circulant matrices and block circulant matrices, a best method is to use various well-known factorizations of circulant matrices and block circulant matrices, and then derive from them general expressions of inverses or Moore-Penrose inverses the matrices. In this chapter we mainly consider Moore-Penrose inverses of block circulant matrices, and then then drive from them some interesting consequences related to sums of matrices. In addition, we shall also consider some extension of the work to quaternion matrices.

For a circulant matrix C over the complex number field $\mathcal{C}$ with the form

$$C = \begin{bmatrix} a_0 & a_1 & \cdots & a_{k-1} \\ a_{k-1} & a_0 & \cdots & a_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_0 \end{bmatrix}, \tag{11.1}$$

the following factorization is well known (see, e.g., Davis [39])

$$U^*CU = \mathrm{diag}(\, \lambda_1, \, \lambda_2, \, \cdots, \, \lambda_k \,), \tag{11.2}$$

where U is a unitary matrix of the form

$$U = (u_{pq})_{k\times k}, \qquad u_{pq} = \frac{1}{\sqrt{k}}\omega^{(p-1)(q-1)}, \qquad \omega^k = 1, \ \ and \ \ \omega \neq 1, \tag{11.3}$$

and

$$\lambda_t = a_0 + a_1\omega^{(t-1)} + a_2(\omega^{(t-1)})^2 + \cdots + a_{k-1}(\omega^{(t-1)})^{k-1}, \quad t = 1, \cdots, k. \tag{11.4}$$

It is evident that the entries in the first row and first column of U are all $1/\sqrt{k}$, and

$$\lambda_1 = a_0 + a_1 + \cdots + a_{k-1}. \tag{11.5}$$

Observe that U in (11.3) is independent of a_0—a_{k-1} in (11.1). Thus (11.2) can directly be extended to block circulant matrix as follows.

Lemma 11.1. *Let*

$$A = \begin{bmatrix} A_1 & A_2 & \cdots & A_k \\ A_k & A_1 & \cdots & A_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_2 & A_3 & \cdots & A_1 \end{bmatrix} \tag{11.6}$$

be a block circulant matrix over the complex number field $\mathcal{C}$, where $A_t \in \mathcal{C}^{m\times n}$, $t = 1, \cdots, k$. Then A satisfies the following factorization equality

$$U_m^* A U_n = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix}, \tag{11.7}$$

where U_r and U_s are two block unitary matrices

$$U_m = (u_{pq}I_m)_{k\times k}, \qquad U_n = (u_{pq}I_n)_{k\times k}, \tag{11.8}$$

u_{pq} is as in (11.3), *meanwhile*

$$J_t = A_1 + A_2\omega^{(t-1)} + A_3(\omega^{(t-1)})^2 + \cdots + A_k(\omega^{(t-1)})^{k-1}, \qquad t = 1, \cdots, k. \tag{11.9}$$

Especially, the block entries in the first block rows and first block columns of U_m and U_n are all scalar products of $1/\sqrt{k}$ with identity matrices, and J_1 is

$$J_1 = A_1 + A_2 + \cdots + A_k. \tag{11.10}$$

Observe that J_1 in (11.7) is the sum of $A_1, A_2, \cdots, A_k$. Thus (11.7) implies that the sum $\sum_{t=1}^k A_t$ is closely linked to its corresponding block circulant matrix through a unitary factorization equality. Recall a fundamental fact in the theory of generalized inverses of matrices (see, e.g., Rao and Mitra [118]) that

$$(PAQ)^\dagger = Q^*A^\dagger P^*, \qquad \text{if } P \text{ and } Q \text{ are unitary}. \tag{11.11}$$

Then from (11.7) we can directly find the following.

Lemma 11.2. *Let A be given in* (11.6), *U_r and U_s be given in* (11.8). *Then the Moore-Penrose inverse of A satisfies*

$$A^\dagger = U_n \mathrm{diag}(\, J_1^\dagger, \; J_2^\dagger, \; \cdots, \; J_k^\dagger \,)U_m^*. \tag{11.12}$$

Proof. Since U_m and U_n in (11.7) are unitary, we find by (11.11) that

$$(U_m^*AU_n)^\dagger = U_n^*A^\dagger U_m.$$

On the other hand, it is easily seen that

$$[\,\mathrm{diag}(\,J_1, \, J_2, \cdots, \, J_k\,)\,]^\dagger = \mathrm{diag}(\,J_1^\dagger, \, J_2^\dagger, \cdots, \, J_k^\dagger\,).$$

Thus (11.12) follows. $\square$

The expression shows that the Moore-Penrose inverse of A can be completely determined by the Moore-Penrose inverses of J_1—J_k. Moreover, $A^\dagger$ is also a block circulant matrix, this fact was pointed out by Cline, Plemmons and Worm in [36] and Smith in [126].

The generalizations of circulants and block circulants have many forms (see, e.g. [29, 36, 39, 144, 145]), and various factorizations of these kinds of matrices can also be established. In that case, one can use the rank formulas in Chapter 8 to those factorizations, and then find from them various expressions for Moore-Penroses inverses of these generalized circulants and generalized block circulants. But we do not intend to go further along this direction. Instead, our next work is to consider some remarkable applications of (11.12) to Moore-Penrose inverses of sums of matrices.

Theorem 11.3(Tian [133, 136]). *Let $A_1, A_2, \cdots, A_k \in \mathcal{C}^{m\times n}$. Then the Moore-Penrose inverse of their sum satisfies*

$$(\,A_1 + A_2 + \cdots + A_k\,)^\dagger = \frac{1}{k}[\,I_n, \, I_n, \cdots, \, I_n\,] \begin{bmatrix} A_1 & A_2 & \cdots & A_k \\ A_k & A_1 & \cdots & A_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_2 & A_3 & \cdots & A_1 \end{bmatrix}^\dagger \begin{bmatrix} I_m \\ I_m \\ \vdots \\ I_m \end{bmatrix}. \tag{11.13}$$

In particular, if the block circulant matrix in it is nonsingular, then

$$(A_1 + A_2 + \cdots + A_k)^{-1} = \frac{1}{k}[I_m, I_m, \cdots, I_m] \begin{bmatrix} A_1 & A_2 & \cdots & A_k \\ A_k & A_1 & \cdots & A_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_2 & A_3 & \cdots & A_1 \end{bmatrix}^{-1} \begin{bmatrix} I_m \\ I_m \\ \vdots \\ I_m \end{bmatrix}. \tag{11.14}$$

Proof. Pre-multiply $[I_n, 0, \cdots, 0]$ and post-multiply $[I_m, 0, \cdots, 0]^T$ on the both sides of (11.12) and observe the structure of U_m and U_n to yield (11.13). □

In [133] and [136], the author proved (11.13) in some direct but tedious methods. New (11.3) is just a simple consequence on the Moore-Penrose inverse of a block circulant matrix. The identity (11.13) manifests that the Moore-Penrose inverse of a sum of matrices can be completely determined through the Moore-Penrose inverse of the corresponding block circulant matrix. In this case, if we can find the expression of the Moore-Penrose inverse of the block circulant matrix by some other methods (not by (11.2)), then we can get the expression for the Moore-Penrose inverse of the sum of matrices. In fact, we have presented many results in Chapter 9 for Moore-Penrose inverses of block matrices. Applying some of them to the block circulant matrix in (11.13), one can derive many new conclusions on Moore-Penrose inverses of sums of matrices. Here we present some of them.

Let A and B be two $m \times n$ matrices. Then according to (11.13) we have

$$(A + B)^{\dagger} = \frac{1}{2}[I_n, \ I_n] \begin{bmatrix} A & B \\ B & A \end{bmatrix}^{\dagger} \begin{bmatrix} I_m \\ I_m \end{bmatrix}. \tag{11.15}$$

As a special case of (11.15), if we replace $A + B$ in (11.15) by a complex matrix $A + iB$, where both A and B are real matrices, then (11.15) becomes the equality

$$(A + iB)^{\dagger} = \frac{1}{2}[I_n, \ I_n] \begin{bmatrix} A & iB \\ iB & A \end{bmatrix}^{\dagger} \begin{bmatrix} I_m \\ I_m \end{bmatrix} = \frac{1}{2}[I_n, \ iI_n] \begin{bmatrix} A & -B \\ B & A \end{bmatrix}^{\dagger} \begin{bmatrix} I_m \\ -iI_m \end{bmatrix}. \tag{11.16}$$

Now applying Theorems 9.8 and 9.9 to (11.15) and (11.16) we find the following two results, which was presented by the author in [136].

Theorem 11.4. *Let A and B be two $m \times n$ complex matrices, and suppose that they satisfy the rank additivity condition*

$$r \begin{bmatrix} A & B \\ B & A \end{bmatrix} = r \begin{bmatrix} A \\ B \end{bmatrix} + r \begin{bmatrix} B \\ A \end{bmatrix} = r[A, \ B] + r[B, \ A], \tag{11.17}$$

or alternatively

$$R(A) \subseteq R(A \pm B) \quad \text{and} \quad R(A^*) \subseteq R(A^* \pm B^*). \tag{11.18}$$

Then

(a) *The Moore-Penrose inverse of $A + B$ can be expressed as*

$$(A + B)^{\dagger} = J^{\dagger}(A) + J^{\dagger}(B) = (E_{B_2} S_A F_{B_1})^{\dagger} + (E_{A_2} S_B F_{A_1})^{\dagger}, \tag{11.19}$$

where $J(A)$ and $J(B)$ are, respectively, the rank complements of A and B in $\begin{bmatrix} A & B \\ B & A \end{bmatrix}$, and

$$S_A = A - BA^{\dagger}B, \quad S_B = B - AB^{\dagger}A, \quad A_1 = E_B A, \quad A_2 = AF_B, \quad B_1 = E_A B, \quad B_2 = BF_A.$$

(b) *The matrices A, B, and the two terms $G_1 = J^{\dagger}(A)$ and $G_2 = J^{\dagger}(B)$ in the right-hand side of* (11.19) *satisfy the following several equalities*

$$r(G_1) = r(A), \qquad r(G_2) = r(B),$$

$$(A+B)(A+B)^{\dagger} = AG_1 + BG_2, \qquad (A+B)^{\dagger}(A+B) = G_1A + G_2B,$$
$$AG_2 + BG_1 = 0, \qquad G_2A + G_1B = 0.$$

Proof. The equivalence of (11.17) and (1.18) is derived from (1.13). We know from Theorem 9.8 that under the condition (11.17), the Moore-Penrose inverse of $\begin{bmatrix} A & B \\ B & A \end{bmatrix}$ can be expressed as

$$\begin{bmatrix} A & B \\ B & A \end{bmatrix}^{\dagger} = \begin{bmatrix} J^{\dagger}(A) & J^{\dagger}(B) \\ J^{\dagger}(B) & J^{\dagger}(A) \end{bmatrix} = \begin{bmatrix} (E_{B_2}S_AF_{B_1})^{\dagger} & (E_{A_2}S_BF_{A_1})^{\dagger} \\ (E_{A_2}S_BF_{A_1})^{\dagger} & (E_{B_2}S_AF_{B_1})^{\dagger} \end{bmatrix}.$$

Then putting it in (11.15) immediately yields (11.19). The results in Part (b) are derived from Theorem 9.9. □

Theorem 11.5. *Let $A + iB$ be an $m \times n$ complex matrix, where A and B are two real matrices, and suppose that A and B satisfy*

$$r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} = r\begin{bmatrix} A \\ B \end{bmatrix} + r\begin{bmatrix} -B \\ A \end{bmatrix} = r[\,A,\ -B\,] + r[\,B,\ A\,], \tag{11.20}$$

or equivalently

$$R(A) \subseteq R(A \pm iB) \quad and \quad R(A^*) \subseteq R(A^T \pm iB^T). \tag{11.21}$$

Then the Moore-Penrose inverse of $A + iB$ can be expressed as

$$(A+iB)^{\dagger} = G_1 - iG_2 = [\,E_{B_2}(A + BA^{\dagger}B)F_{B_1}\,]^{\dagger} - i[\,E_{A_2}(B + AB^{\dagger}A)F_{A_1}\,]^{\dagger}, \tag{11.22}$$

where $A_1 = E_BA,\ A_2 = AF_B,\ B_1 = E_AB$ and $B_2 = BF_A$.

Proof. Follows directly from Theorem 11.4. □

Corollary 11.6. *Suppose that $A + iB$ is a nonsingular complex matrix, where A and B are real.*

(a) *If both A and B are nonsingular, then*

$$(A+iB)^{-1} = (A + BA^{-1}B)^{-1} - i(B + AB^{-1}A)^{-1}.$$

(b) *If both $R(A) \cap R(B) = \{0\}$ and $R(A^*) \cap R(B^*) = \{0\}$, then*

$$(A+iB)^{-1} = (E_BAF_B)^{\dagger} - i(E_ABF_A)^{\dagger}.$$

(c) *Let $A = \lambda I_m$, where λ is a real number such that $\lambda I_m + iB$ is nonsingular, then*

$$(\lambda I_m + iB)^{-1} = \lambda(\lambda^2 I_m + B^2)^{-1} - i(\lambda^2 B + B^{\dagger}B^3B^{\dagger})^{\dagger}.$$

Proof. Follows directly from Theorem 11.4. □

As a special case of Theorem 11.5, we have the following interesting result: Suppose $M = A + iB$ is a nipotent matrix, i.e., $M^2 = 0$. Then its Moore-Penrose inverse can be expressed as

$$(A+iB)^{\dagger} = (E_BAF_B)^{\dagger} - i(E_ABF_A)^{\dagger}.$$

we leave it as an exercise to the reader.

(b) If both $R(A) \cap R(B) = \{0\}$ and $R(A^*) \cap R(B^*) = \{0\}$, then

$$(A+iB)^{-1} = (E_BAF_B)^{\dagger} - i(E_ABF_A)^{\dagger}.$$

We next turn our attention to the Moore-Penrose inverse of sum of k matrices, and give some general formulas.

Theorem 11.7. *Let $A_1, A_2, \cdots, A_k \in \mathcal{C}^{m\times n}$ be given. If they satisfy the following rank additivity condition*

$$r(A) = kr[\,A_1,\ \cdots,\ A_k\,] = kr[\,A_1^*,\ \cdots,\ A_k^*\,], \tag{11.23}$$

where A is the circulant block matrix defined in (11.6)*, then*

(a) *The Moore-Penrose inverse of the sum $\sum_{i=1}^{k} A_i$ can be expressed as*

$$(A_1 + A_2 + \cdots + A_k)^{\dagger} = J^{\dagger}(A_1) + J^{\dagger}(A_2) + \cdots + J^{\dagger}(A_k), \tag{11.24}$$

where $J(A_i)$ is the rank complement of $A_i (1 \le i \le k)$ in A.

(b) *The rank of $J(A_i)$ is*

$$r[J(A_i)] = r[\, A_1, \, \cdots, \, A_k \,] + r[\, A_1^*, \, \cdots, \, A_k^* \,] - r(A) + r(D_i), \tag{11.25}$$

where $1 \le i \le k$, D_i is the $(k-1) \times (k-1)$ block matrix resulting from the deletion of the first block row and ith block column of A.

(c) *$A_1, \, A_2, \, \cdots, \, A_k$ and $J^{\dagger}(A_1), \, J^{\dagger}(A_2), \, \cdots, \, J^{\dagger}(A_k)$ satisfy the following two equalities*

$$(A_1 + \cdots + A_k)(A_1 + \cdots + A_k)^{\dagger} = A_1 J^{\dagger}(A_1) + \cdots + A_k J^{\dagger}(A_k),$$
$$(A_1 + \cdots + A_k)^{\dagger}(A_1 + \cdots + A_k) = J^{\dagger}(A_1) A_1 + \cdots + J^{\dagger}(A_k) A_k.$$

Proof. Follows from combining Theorem 9.18 with the equality (11.13). □

Corollary 11.8. *Let $A_1, \, A_2, \, \cdots, \, A_k \in \mathcal{C}^{m \times n}$. If they satisfy the following rank additivity condition*

$$r(A_1 + A_2 + \cdots + A_k) = r(A_1) + r(A_2) + \cdots + r(A_k), \tag{11.26}$$

then the Moore-Penrose inverse of the sum $\sum_{i=1}^{k} A_i$ can be expressed as

$$(A_1 + A_2 + \cdots + A_k)^{\dagger} = (E_{\alpha_1} A_1 F_{\beta_1})^{\dagger} + (E_{\alpha_2} A_2 F_{\beta_2})^{\dagger} + \cdots + (E_{\alpha_k} A_k F_{\beta_k})^{\dagger}, \tag{11.27}$$

where α_i and β_i are

$$\alpha_i = [\, A_1, \, \cdots \, A_{i-1}, \, A_{i+1}, \, \cdots, \, A_k \,], \quad \beta_i = \begin{bmatrix} A_1 \\ \vdots \\ A_{i-1} \\ A_{i+1} \\ \vdots \\ A_k \end{bmatrix}, \quad i = 1, \, 2, \, \cdots, \, k.$$

Proof. We first show that under the condition (11.17) the rank of the circulant matrix A in (11.6) is

$$r(A) = k[\, r(A_1) + r(A_2) + \cdots + r(A_k) \,]. \tag{11.28}$$

According to (11.7), we see that

$$r(A) = r(J_1) + r(J_2) + \cdots + r(J_k).$$

Under Eq. (11.26), the ranks of all J_i are the same, that is,

$$r(J_i) = r(A_1) + r(A_2) + \cdots + r(A_k), \qquad i = 1, \, 2, \, \cdots, \, k.$$

Thus we have (11.28). In that case, applying the result in Corollary 9.19 to the circulant block matrix A in (11.13) produces the equality (11.27). □

It is worth to point out that the formulas for Moore-Penrose inverses of sums of matrices given in this chapter and those for Moore-Penrose inverses of block matrices given in Chapter 9 are, in fact, a group of dual results. That is to say, not only can we derive Moore-Penrose inverses of sums of matrices from Moore-Penrose inverses of block matrices, but also we can make a contrary derivation. For simplicity, here we only illustrate this assertion by a 2×2 block matrix. In fact, for any 2×2 block matrix can factor as

$$M = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} + \begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix} = N_1 + N_2.$$

If M satisfies the rank additivity condition (9.22), then N_1 and N_2 satisfy

$$r\begin{bmatrix} N_1 & N_2 \\ N_2 & N_1 \end{bmatrix} = r\begin{bmatrix} A & 0 & 0 & B \\ 0 & D & C & 0 \\ 0 & B & A & 0 \\ C & 0 & 0 & D \end{bmatrix} = 2r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = 2r\begin{bmatrix} N_1 \\ N_2 \end{bmatrix} = 2r[\, N_1,\ N_2\,].$$

Hence by Theorem 11.4, we have

$$M_1^\dagger = (\, N_1 + N_2\,)^\dagger = J^\dagger(N_1) + J^\dagger(N_2), \tag{11.29}$$

where $J(N_1)$ and $J(N_2)$ are, respectively, the rank complements of N_1 and N_2 in $\begin{bmatrix} N_1 & N_2 \\ N_2 & N_1 \end{bmatrix}$. Written in an explicit form, (11.29) is exactly the formula (9.33).

Besides (11.13), some other identities between Moore-Penrose inverses of sums of matrices and Moore-Penrose inverses of block matrices can also be established. Here we present a result for the sum of four matrices.

$$(\, A_0 + A_1 + A_2 + A_3\,)^\dagger = \frac{1}{4}[\, I_n,\, I_n,\, I_n,\, I_n\,]\begin{bmatrix} A_0 & A_1 & A_2 & A_3 \\ A_1 & A_0 & A_3 & A_2 \\ A_2 & A_3 & A_0 & A_1 \\ A_3 & A_2 & A_1 & A_0 \end{bmatrix}^\dagger \begin{bmatrix} I_m \\ I_m \\ I_m \\ I_m \end{bmatrix}, \tag{11.30}$$

Clearly, the block matrix in (9.30) is not 4×4 block circulant, but 2×2 block circulant with two 2×2 block circulants in it.

In addition, we mention another interesting fact that (11.16) can be extended to any real quaternion matrix of the form $A = A_0 + iA_1 + jA_2 + kA_3$, where A_0—A_3 are real $m\times n$ matrices and $i^2 = j^2 = k^2 = -1$, $ijk = -1$, as follows:

$$(\, A_0 + iA_1 + iA_2 + kA_3\,)^\dagger = \frac{1}{2}[\, I_n,\ jI_n\,]\begin{bmatrix} A_0 + iA_1 & -(\, A_2 + iA_3\,) \\ A_2 - iA_3 & A_0 - iA_1 \end{bmatrix}^\dagger \begin{bmatrix} I_m \\ -jI_m \end{bmatrix}, \tag{11.31}$$

and

$$(\, A_0 + iA_1 + iA_2 + kA_3\,)^\dagger = \frac{1}{4}[\, I_n,\ iI_n,\ jI_n,\ kI_n\,]\begin{bmatrix} A_0 & -A_1 & -A_2 & -A_3 \\ A_1 & A_0 & A_3 & -A_2 \\ A_2 & -A_3 & A_0 & A_1 \\ A_3 & A_2 & -A_1 & A_0 \end{bmatrix}^\dagger \begin{bmatrix} I_m \\ -iI_m \\ -jI_m \\ -kI_m \end{bmatrix}. \tag{11.32}$$

Moreover denote $(\, A_0 + iA_1 + iA_2 + kA_3\,)^\dagger = G_0 + iG_1 + iG_2 + kG_3$. Then

$$\begin{bmatrix} A_0 & -A_1 & -A_2 & -A_3 \\ A_1 & A_0 & A_3 & -A_2 \\ A_2 & -A_3 & A_0 & A_1 \\ A_3 & A_2 & -A_1 & A_0 \end{bmatrix}^\dagger = \begin{bmatrix} G_0 & -G_1 & -G_2 & -G_3 \\ G_1 & G_0 & G_3 & -G_2 \\ G_2 & -G_3 & G_0 & G_1 \\ G_3 & G_2 & -G_1 & G_0 \end{bmatrix}. \tag{11.33}$$

These equalities are in fact derived from the following two universal factorization equalities (see [137])

$$P_{2m}\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} P_{2n}^* = \begin{bmatrix} A_0 + A_1 i & -(A_2 + A_3 i) \\ A_2 - A_3 i & A_0 - A_1 i \end{bmatrix}, \tag{11.34}$$

where P_{2m} and P_{2n}^* are the following two unitary quaternion matrices

$$P_{2m} = \frac{1}{\sqrt{2}}\begin{bmatrix} I_m & -iI_m \\ -jI_m & kI_m \end{bmatrix}, \qquad P_{2n}^* = \frac{1}{\sqrt{2}}\begin{bmatrix} I_n & jI_n \\ iI_n & -kI_n \end{bmatrix},$$

and

$$Q_{4m}\begin{bmatrix} A & & & \\ & A & & \\ & & A & \\ & & & A \end{bmatrix} Q_{4n}^* = \begin{bmatrix} A_0 & -A_1 & -A_2 & -A_3 \\ A_1 & A_0 & -A_3 & A_2 \\ A_2 & A_3 & A_0 & -A_1 \\ A_3 & -A_2 & A_1 & A_0 \end{bmatrix}, \tag{11.35}$$

where Q_{4t} is the following unitary quaternion matrix

$$Q_{4t} = Q_{4t}^{*} = \frac{1}{2}\begin{bmatrix} I_t & iI_t & jI_t & kI_t \\ -iI_t & I_t & kI_t & -jI_t \\ -jI_t & -kI_t & I_t & iI_t \\ -kI_t & jI_t & -iI_t & I_t \end{bmatrix}, \quad t = m,\ n.$$

Based on (11.31)—(11.33), one find easily determine expressions of Moore-Penrose inverses of any real quaternion matrices, especially the inverses of nonsingular matrices.

Furthermore, it should be pointed out that the above work can extend to matrices over any 2^n-dimensional real and complex Clifford algebras through a set of universal similarity factorization equalities established in the author's recent papers [137] and [138].

Chapter 12

Rank equalities for submatrices in Moore-Penrose inverses

Let

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{12.1}$$

be a 2×2 block matrix over $\mathcal{C}$, where $A \in \mathcal{C}^{m \times n}, \ B \in \mathcal{C}^{m \times k}, \ C \in \mathcal{C}^{l \times n}, \ D \in \mathcal{C}^{l \times k}$, and let

$$V_1 = \begin{bmatrix} A \\ C \end{bmatrix}, \qquad V_2 = \begin{bmatrix} B \\ D \end{bmatrix}, \qquad W_1 = [\, A, \ B \,], \qquad W_2 = [\, C, \ D \,]. \tag{12.2}$$

Moreover, partition the Moore-Penrose inverse of M as

$$M^{\dagger} = \begin{bmatrix} G_1 & G_2 \\ G_3 & G_4 \end{bmatrix}, \tag{12.3}$$

where $G_1 \in \mathcal{C}^{n \times m}$. As is well known, the expressions of the submatrices G_1—G_4 in (12.3) are quite complicated if there are no restrictions on the blocks in M (see, e.g., Hung and Markham [73], Miao [97]). In that case, it is hard to find properties of submatrices in $M^{\dagger}$. In the present chapter, we consider a simpler problem—what is the ranks of submatrices in $M^{\dagger}$, when M is arbitrarily given? This problem was examined by Robinson [120] and Tian [134]. In this chapter, we shall give this problem a new discussion.

Theorem 12.1. *Let M and $M^{\dagger}$ be given by* (12.1) *and* (12.3). *Then*

$$r(G_1) = r\begin{bmatrix} V_2D^*W_2 & V_1 \\ W_1 & 0 \end{bmatrix} - r(M), \quad r(G_2) = r\begin{bmatrix} V_2B^*W_1 & V_1 \\ W_2 & 0 \end{bmatrix} - r(M), \tag{12.4}$$

$$r(G_3) = r\begin{bmatrix} V_1C^*W_2 & V_2 \\ W_1 & 0 \end{bmatrix} - r(M), \quad r(G_4) = r\begin{bmatrix} V_1A^*W_1 & V_2 \\ W_2 & 0 \end{bmatrix} - r(M), \tag{12.5}$$

where $V_1, \ V_2, \ W_1$ and W_2 are defined in (12.2).

Proof. We only show the first equality in (12.4). In fact G_1 in (12.3) can be written as

$$G_1 = [\, I_n, \ 0 \,]M^{\dagger}\begin{bmatrix} I_m \\ 0 \end{bmatrix} = PM^{\dagger}Q. \tag{12.6}$$

Then applying (2.1) to it we find

$$\begin{aligned} r(G_1) &= r\begin{bmatrix} M^*MM^* & M^*Q \\ PM^* & 0 \end{bmatrix} - r(M) \\ &= r\begin{bmatrix} MM^*M & MP^* \\ Q^*M & 0 \end{bmatrix} - r(M) \end{aligned}$$

$$= r\begin{bmatrix} [V_1, V_2]M^*\begin{bmatrix} W_1 \\ W_2 \end{bmatrix} & V_1 \\ W_1 & 0 \end{bmatrix} - r(M)$$
$$= r\begin{bmatrix} [0, V_2]M^*\begin{bmatrix} 0 \\ W_2 \end{bmatrix} & V_1 \\ W_1 & 0 \end{bmatrix} - r(M) = r\begin{bmatrix} V_2D^*W_2 & V_1 \\ W_1 & 0 \end{bmatrix} - r(M),$$

establishing the first equality in (12.4). □

Corollary 12.2. *Let M and $M^\dagger$ be given by* (12.1) *and* (12.3). *If*

$$r(M) = r(V_1) + r(V_2), \quad i.e., \quad R(V_1) \cap R(V_2) = \{0\}, \tag{12.7}$$

then

$$r(G_1) = r\begin{bmatrix} A & B \\ D^*C & D^*D \end{bmatrix} - r\begin{bmatrix} B \\ D \end{bmatrix}, \quad r(G_2) = r\begin{bmatrix} B^*A & B^*B \\ C & D \end{bmatrix} - r\begin{bmatrix} B \\ D \end{bmatrix}, \tag{12.8}$$

$$r(G_3) = r\begin{bmatrix} A & B \\ C^*C & C^*D \end{bmatrix} - r\begin{bmatrix} A \\ C \end{bmatrix}, \quad r(G_4) = r\begin{bmatrix} A^*A & A^*B \\ C & D \end{bmatrix} - r\begin{bmatrix} A \\ C \end{bmatrix}. \tag{12.9}$$

Proof. Under (12.7), we also know that $R(V_1) \cap R(V_2D^*W_2) = \{0\}$. Thus the first equality in (12.4) becomes

$$r(G_1) = r\begin{bmatrix} V_2D^*W_2 & V_1 \\ W_1 & 0 \end{bmatrix} - r(M)$$
$$= r\begin{bmatrix} V_2D^*W_2 \\ W_1 \end{bmatrix} + r(V_1) - r(M) = r\begin{bmatrix} D^*W_2 \\ W_1 \end{bmatrix} - r(V_2) = r\begin{bmatrix} W_1 \\ D^*W_2 \end{bmatrix} - r(V_2),$$

establishing the first one in (12.8). Similarly, we can show the other three in (12.8) and (12.9). □

Similarly, we have the following.

Corollary 12.3. *Let M and $M^\dagger$ be given by* (12.1) *and* (12.3). *If*

$$r(M) = r(W_1) + r(W_2), \quad i.e., \quad R(W_1^*) \cap R(W_2^*) = \{0\}, \tag{12.10}$$

then

$$r(G_1) = r\begin{bmatrix} A & BD^* \\ C & DD^* \end{bmatrix} - r[C, D], \quad r(G_2) = r\begin{bmatrix} A & BB^* \\ C & DB^* \end{bmatrix} - r[A, B], \tag{12.11}$$

$$r(G_3) = r\begin{bmatrix} AC^* & B \\ CC^* & D \end{bmatrix} - r[C, D], \quad r(G_4) = r\begin{bmatrix} AA^* & B \\ CA^* & D \end{bmatrix} - r[A, B]. \tag{12.12}$$

Combining the above two corollaries, we obtain the following, which is previously shown in Corollary 9.9.

Corollary 12.4. *Let M and $M^\dagger$ be given by* (12.1) *and* (12.3). *If M satisfies the rank additivity condition*

$$r(M) = r(V_1) + r(V_2) = r(W_1) + r(W_2), \tag{12.13}$$

then

$$r(G_1) = r(D) + r(V_1) + r(W_1) - r(M), \tag{12.14}$$
$$r(G_2) = r(B) + r(V_1) + r(W_2) - r(M), \tag{12.15}$$
$$r(G_3) = r(C) + r(V_2) + r(W_1) - r(M), \tag{12.16}$$
$$r(G_4) = r(A) + r(V_2) + r(W_2) - r(M). \tag{12.17}$$

Proof. We only show (12.14). Under (12.13), we find that

$$r\begin{bmatrix} V_2D^*W_2 & V_1 \\ W_1 & 0 \end{bmatrix} = r(V_2D^*W_2) + r(V_1) + r(W_1),$$

where

$$r(V_2D^*W_2) = r\begin{bmatrix} BD^*C & BD^*D \\ DD^*C & DD^*D \end{bmatrix} = r(D).$$

Thus the first equality in (12.4) reduces to (12.14). $\square$

Corollary 12.5. *Let M and $M^\dagger$ be given by* (12.1) *and* (12.3). *If M satisfies the rank additivity condition*

$$r(M) = r(A) + r(B) + r(C) + r(D), \tag{12.18}$$

then

$$r(G_1) = r(A), \quad r(G_2) = r(C), \quad r(G_3) = r(B), \quad r(G_4) = r(D). \tag{12.19}$$

Proof. Follows directly from (12.14)—(12.17). $\square$

Corollary 12.6. *Let M and $M^\dagger$ be given by* (12.1) *and* (12.3). *If*

$$r(M) = r(V_1), \quad i.e., \quad R(V_2) \subseteq R(V_1), \tag{12.20}$$

then

$$r(G_1) = r(A), \qquad r(G_2) = r(C), \tag{12.21}$$

$$r(G_3) = r\begin{bmatrix} V_1C^*C & V_2 \\ A & 0 \end{bmatrix} - r(V_1), \quad r(G_4) = r\begin{bmatrix} V_1A^*A & V_2 \\ C & 0 \end{bmatrix} - r(V_1). \tag{12.22}$$

Proof. The inclusion in (12.20) implies that

$$R(V_2D^*W_2) \subseteq R(V_1), \quad R(V_2D^*W_1) \subseteq R(V_1), \quad R(B) \subseteq R(A), \quad R(D) \subseteq R(C).$$

Thus the two rank equalities in (12.4) become

$$r(G_1) = r\begin{bmatrix} V_2D^*W_2 & V_1 \\ W_1 & 0 \end{bmatrix} - r(M) = r(V_1) + r(W_1) - r(M) = r(W_1) = r(A),$$

$$r(G_2) = r\begin{bmatrix} V_2B^*W_1 & V_1 \\ W_2 & 0 \end{bmatrix} - r(M) = r(V_1) + r(W_2) - r(M) = r(W_2) = r(C),$$

and the two rank equalities in (12.5) become

$$\begin{aligned} r(G_3) &= r\begin{bmatrix} V_1C^*W_2 & V_2 \\ W_1 & 0 \end{bmatrix} - r(M) \\ &= r\begin{bmatrix} V_1C^*C & V_1C^*D & V_2 \\ A & B & 0 \end{bmatrix} - r(M) = r\begin{bmatrix} V_1C^*C & V_2 \\ A & 0 \end{bmatrix} - r(M), \end{aligned}$$

$$\begin{aligned} r(G_4) &= r\begin{bmatrix} V_1A^*W_1 & V_2 \\ W_2 & 0 \end{bmatrix} - r(M) \\ &= r\begin{bmatrix} V_1A^*A & V_1A^*B & V_2 \\ C & D & 0 \end{bmatrix} - r(M) = r\begin{bmatrix} V_1A^*A & V_2 \\ C & 0 \end{bmatrix} - r(M). \end{aligned}$$

Hence we have (12.21) and (12.22). $\square$

Similarly, we have the following.

Corollary 12.7. *Let M and $M^\dagger$ be given by* (12.1) *and* (12.3). *If*

$$r(M) = r(W_1), \quad i.e., \quad R(W_2^*) \subseteq R(W_1^*), \tag{12.23}$$

then

$$r(G_1) = r(A), \qquad r(G_3) = r(B), \tag{12.24}$$

$$r(G_2) = r\begin{bmatrix} BB^*W_1 & A \\ W_2 & 0 \end{bmatrix} - r(W_1), \quad r(G_4) = r\begin{bmatrix} AA^*W_1 & B \\ W_2 & 0 \end{bmatrix} - r(W_1). \tag{12.25}$$

Combining the above two corollaries, we obtain the following.

Corollary 12.8. *Let M and $M^\dagger$ be given by* (12.1) *and* (12.3). *If*

$$r(M) = r(A), \tag{12.26}$$

then

$$r(G_1) = r(A), \qquad r(G_2) = r(C), \qquad r(G_3) = r(B), \tag{12.27}$$

and

$$r(G_4) = r\begin{bmatrix} AA^*A & B \\ C & 0 \end{bmatrix} - r(A). \tag{12.28}$$

Proof. Clearly (12.26) implies that $r(M) = r(V_1) = r(W_1)$. Thus we have (12.27) by Corollaries 12.6 and 12.7. On the other hand, (12.26) is also equivalent to $AA^\dagger B = B$, $CA^\dagger A = C$ and $D = CA^\dagger B$ by (1.5). Hence

$$\begin{aligned} r(G_4) &= r\begin{bmatrix} V_1A^*W_1 & V_2 \\ W_2 & 0 \end{bmatrix} - r(M) \\ &= r\begin{bmatrix} AA^*W_1 & B \\ CA^*W_1 & D \\ W_2 & 0 \end{bmatrix} - r(A) \\ &= r\begin{bmatrix} AA^*W_1 & B \\ W_2 & 0 \end{bmatrix} - r(W_1) \\ &= r\begin{bmatrix} AA^*A & AA^*B & B \\ C & D & 0 \end{bmatrix} - r(A) = r\begin{bmatrix} AA^*A & B \\ C & 0 \end{bmatrix} - r(A), \end{aligned}$$

which is (12.28). $\square$

Next we list a group of rank inequalities derived from (12.4) and (12.5).

Corollary 12.9. *Let M and $M^\dagger$ be given by* (12.1) *and* (12.2). *Then the rank of G_1 in $M^\dagger$ satisfies the rank inequalities*

$$r(G_1) \le r(D) + r[\,A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r(M), \tag{12.29}$$

$$r(G_1) \ge r[\,A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r(M), \tag{12.30}$$

$$r(G_1) \ge r(D) - r[\,C,\ D\,] - r\begin{bmatrix} B \\ D \end{bmatrix} + r(M). \tag{12.31}$$

Proof. Observe that

$$r(V_1) + r(W_1) \le r\begin{bmatrix} V_2D^*W_2 & V_1 \\ W_1 & 0 \end{bmatrix} \le r(D) + r(V_1) + r(W_1).$$

Putting them in the first rank equality in (12.4), we obtain (12.29) and (12.30). To show (12.31), we need the following rank equality

$$r(CA^\dagger B) \ge r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r\begin{bmatrix} A \\ C \end{bmatrix} - r[\,A,\ B\,] + r(A), \tag{12.32}$$

which is derived from (1.6). Now applying (12.32) to $PM^\dagger Q$ in (12.6), we obtain

$$\begin{aligned} r(G_1) = r(PM^\dagger Q) &\ge r\begin{bmatrix} M & Q \\ P & 0 \end{bmatrix} - r\begin{bmatrix} M \\ P \end{bmatrix} - r[\,M,\ Q\,] + r(M) \\ &= r(D) - r[\,C,\ D\,] - r\begin{bmatrix} B \\ D \end{bmatrix} + r(M), \end{aligned}$$

which is (12.31). $\square$

Rank inequalities for the block entries $G_2,\ G_3$ and G_4 in (12.3) can also be derived in the similar way shown above. Finally let $D = 0$ in (12.1). Then the results in (12.4) and (12.5) can be simplified to the following.

Thoerem 12.10. *Let*

$$M_1 = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}, \tag{12.33}$$

and denote the Moore-Penrose inverse of M as

$$M_1^{\dagger} = \begin{bmatrix} G_1 & G_2 \\ G_3 & G_4 \end{bmatrix}, \tag{12.34}$$

where $G_1 \in \mathcal{C}^{n \times m}$. Then

$$r(G_1) = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A,\ B\,] - r(M_1), \qquad r(G_2) = r(C), \qquad r(G_3) = r(B), \tag{12.35}$$

$$r(G_4) = r\begin{bmatrix} AA^*A & AA^*B & B \\ CA^*A & CA^*B & 0 \\ C & 0 & 0 \end{bmatrix} - r(M_1). \tag{12.36}$$

Various consequences of (12.35) and (12.36) can also be derived. But we omit them here.

Chapter 13

Rank equalities for Drazin inverses

As one of the important types of generalized inverses of matrices, the Drazin inverses and their applications have well been examined in the literature. Having established so many rank equalities in the preceding chapters, one might naturally consider how to extend that work from Moore-Penrose inverses to Drazin inverses. To do this, we only need to use a basic identity on the Drazin inverse of a matrix $A^D = A^k(A^{2k+1})^{\dagger}A^k$ (see, e.g., Campbell and Meyer [21]). In that case, the rank formulas obtained in the preceding chapters can all apply to establish various rank equalities for matrix expressions involving Drazin inverses of matrices.

Theorem 13.1. *Let* $A \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$. *Then*

(a) $r(\, I_m \pm A^D\,) = r(\, I_m \pm A\,)$.

(b) $r[\, I_m - (A^D)^2\,] = r(\, I_m - A^2\,)$.

Proof. Observe that $R(A^k) = R(A^{2k+1})$ and $R[(A^k)^*] = R[(A^{2k+1})^*]$. Thus applying (1.7) and then (1.16) to $I_m - A^D = I_m - A^k(A^{2k+1})^{\dagger}A^k$ yields

$$\begin{aligned} r(\, I_m - A^D\,) &= r[\, I_m - A^k(A^{2k+1})^{\dagger}A^k\,] \\ &= r\begin{bmatrix} A^{2k+1} & A^k \\ A^k & I_m \end{bmatrix} - r(A^{2k+1}) \\ &= r\begin{bmatrix} A^{2k+1} - A^{2k} & 0 \\ 0 & I_m \end{bmatrix} - r(A^k) \\ &= r(\, A^{2k+1} - A^{2k}\,) + m - r(A^k) = r(A^{2k}) + r(\, I_m - A\,) - r(A^k) = r(\, I_m - A\,). \end{aligned}$$

Similarly we can find $r(\, I_m + A^D\,) = r(\, I_m + A\,)$. Note by (1.12) that

$$r[\, I_m - (A^D)^2\,] = r(\, I_m + A^D\,) + r(\, I_m - A^D\,) - m = r(\, I_m + A\,) + r(\, I_m - A\,) - m = r(\, I_m - A^2\,),$$

establishing Part (b). $\square$

Theorem 13.2. *Let* $A \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$. *Then*

(a) $r(\, A - AA^D\,) = r(\, A - A^DA\,) = r(\, A - A^2\,)$.

(b) $r(\, A - AA^DA\,) = r(A) - r(A^D)$, *i.e.*, $AA^DA \leq_{rs} A$.

(c) $AA^D = A^DA = A \Leftrightarrow A^2 = A$.

The results in Theorem 13.2(d) is well known, see, e.g., Campbell and Meyer [21].

Proof. Applying (1.6) and (1.16) to $A - AA^D$ yields

$$\begin{aligned} r(\, A - AA^D\,) &= r[\, A - A^{k+1}(A^{2k+1})^{\dagger}A^k\,] \\ &= r\begin{bmatrix} A^{2k+1} & A^{k+1} \\ A^k & A \end{bmatrix} - r(A^{2k+1}) \\ &= r\begin{bmatrix} A^{2k+1} - A^{2k} & 0 \\ 0 & A \end{bmatrix} - r(A^k) \end{aligned}$$

$$
\begin{aligned}
&= r(A^{2k+1} - A^{2k}) + r(A) - r(A^k) \\
&= r(A^{2k}) + r(I_m - A) - m + r(A) - r(A^k). \\
&= r(A - A^2).
\end{aligned}
$$

as required for Part (a). Notice that A^D is an outer inverse of A. Thus it follow by (5.6) that

$$r(A - AA^DA) = r(A) - r(A^D) = r(A) - r(A^k),$$

as required for Part (b). The results in Parts (c) and (d) follow from Part (a). □

Theorem 13.3. *Let* $A \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$.
(a) $r(A - A^D) = r(A^{k+2} - A^k) + r(A) - r(A^k) = r(A - A^3)$.
(b) $r(A - A^D) = r(A) - r(A^D)$, *i.e.*, $A^D \leq_{rs} A \Leftrightarrow A^{k+2} = A^k$.
(c) $r(A - A^{\#}) = r(A^3 - A)$, *if* $\mathrm{Ind}(A) = 1$.
(d) $A^{\#} = A \Leftrightarrow A^3 = A$.

Proof. Applying (1.6) and (1.16) to $A - A^D$ yields

$$
\begin{aligned}
r(A - A^D) &= r[A - A^k(A^{2k+1})^{\dagger}A^k] \\
&= r\begin{bmatrix} A^{2k+1} & A^k \\ A^k & A \end{bmatrix} - r(A^{2k+1}) \\
&= r\begin{bmatrix} A^{2k+1} - A^{2k-1} & 0 \\ 0 & A \end{bmatrix} - r(A^k) \\
&= r(A^{k+2} - A^k) + r(A) - r(A^k) \\
&= r(A^k) + r(I_m - A^2) - m + r(A) - r(A^k) \\
&= r(A - A^3),
\end{aligned}
$$

as required for Part(a). The results in Parts (b), (c) and (d) follow immediately from it, where the result in Part (d) is well known. . □

Similarly, we can establish the following two.

Theorem 13.4. *Let* $A \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$.
(a) *If* $k \geq 2$, *then* $r(A^D - A^2) = r(A^5 - A^2)$.
(b) *If* $k = 1$, *then* $r(A^{\#} - A^2) = r(A^4 - A)$.
(c) $A^2 = A^D \Leftrightarrow A^5 = A^2$.
(d) $A^2 = A^{\#} \Leftrightarrow A^4 = A$ *when* $k = 1$.

The two equivalence relations in Theorem 13.4(c) and (d) were obtained by Grass and Trenkler [55] when they considered generalized and hypergeneralized projectors.

Theorem 13.5. *Let* $A \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$.
(a) $r(A^D - A^3) = r(A^7 - A^3)$.
(b) *If* $k = 2$, *then* $r(A^D - A^3) = r(A^6 - A^2)$.
(c) *If* $k = 1$, *then* $r(A^{\#} - A^3) = r(A^5 - A)$.
(d) $A^3 = A^D \Leftrightarrow A^7 = A^3$.
(e) $A^3 = A^D \Leftrightarrow A^6 = A^2$ *when* $k = 2$.
(f) $A^3 = A^{\#} \Leftrightarrow A^5 = A$ *when* $k = 1$.
(g) *In general*, $r(A^D - A^t) = r(A^{2t+1} - A^t)$.

Theorem 13.6. *Let* $A \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$.
(a) $r(A^D - AA^DA) = r(A^k - A^{k+2})$.
(b) $A^D = AA^DA \Leftrightarrow A^{k+2} = A^k$.

Proof. Observe that

$$A^{k+1}(A^D - AA^DA) = A^k - A^{k+2} \quad \text{and} \quad (A^D)^{k+1}(A^k - A^{k+2}) = A^D - AA^DA.$$

Thus Part (a) follows. $\square$

A square matrix A is said to be quasi-idempotent if $A^{k+1} = A^k$ for some positive integer k. In a recent paper by Mitra [104], quasi-idempotent matrices and the related topics are well examined. The result given below reveals a new aspect on quasi-idempotent matrix.

Theorem 13.7. *Let $A \in \mathcal{C}^{m\times m}$ with* $\mathrm{Ind}(A) = k$. *Then*
(a) $r[\, A^D \pm (A^D)^2\,] = r(\, A^{k+1} \pm A^k\,)$.
(b) $(A^D)^2 = A^D$, *i.e., A^D is idempotent* $\Leftrightarrow A^{k+1} = A^k$, *i.e., A is quasi-idempotent.*
(c) $r[\, A^{\#} \pm (A^{\#})^2\,] = r(\, A^2 \pm A\,)$.
(d) *$A^{\#}$ is idempotent $\Leftrightarrow$ A is idempotent.*

Proof. By (2.3) and $(A^D)^2 = (A^2)^D$ we find that

$$\begin{aligned} & r[\, A^D - (A^D)^2\,] \\ = \; & r[\, A^k(A^{2k+1})^{\dagger}A^k - A^{2k}(A^{4k+2})^{\dagger}A^{2k}\,] \\ = \; & r\begin{bmatrix} -A^{2k+1} & 0 & A^k \\ 0 & A^{4k+2} & A^{2k} \\ A^k & A^{2k} & 0 \end{bmatrix} - r(A^{2k+1}) - r(A^{4k+2}) \\ = \; & r\begin{bmatrix} -A^{2k+1} & 0 & A^k \\ 0 & A^{2k+2} & A^k \\ A^k & A^k & 0 \end{bmatrix} - 2r(A^k) \\ = \; & r\begin{bmatrix} 0 & 0 & A^k \\ 0 & A^{2k+2} - A^{2k+1} & 0 \\ A^k & 0 & 0 \end{bmatrix} - 2r(A^k) = r(\, A^{2k+2} - A^{2k+1}\,) = r(\, A^{k+1} - A^k\,). \end{aligned}$$

Similarly, we can obtain $r[\, A^D + (A^D)^2\,] = r(\, A^{k+1} + A^k\,)$. The results in Parts (b)—(d) follow immediately from Part (a). $\square$

Theorem 13.8. *Let $A \in \mathcal{C}^{m\times m}$ with* $\mathrm{Ind}(A) = k$. *Then*
(a) $r[\, A^D - (A^D)^3\,] = r(\, A^k - A^{k+2}\,)$.
(b) $r[\, A^{\#} - (A^{\#})^3\,] = r(\, A - A^3\,)$, *if* $\mathrm{Ind}(A) = 1$.
(c) $(A^D)^3 = A^D \Leftrightarrow A^D = AA^DA \Leftrightarrow A^{k+2} = A^k$.
(d) *$A^{\#}$ is tripotent if and only if A is tripotent.*

Proof. Notice that

$$A[\, A^D - (A^D)^3\,]A = AA^DA - A^D \quad \text{and} \quad A^D(\, AA^DA - A^D\,)A^D = A^D - (A^D)^3.$$

Thus we have $r[\, A^D - (A^D)^3\,] = r(\, AA^DA - A^D\,)$. In that case we have Part (a) by Theorem 13.6(a). yields

$$\begin{aligned} r[\, A^D - (A^D)^3\,] &= r[\, A^D + (A^D)^2\,] + r[\, A^D - (A^D)^2\,] - r(A^D) \\ &= r(\, A^{k+1} + A^k\,) + r(\, A^{k+1} - A^k\,) - r(A^k) = r(\, A^{k+1} + A^k\,) + r(\, A^{k+1} - A^k\,) - r(A^k), \end{aligned}$$

as required for Part (a). The equivalence in Part (c) follows directly from Part (a). $\square$

Theorem 13.9. *Let $A \in \mathcal{C}^{m\times m}$ with* $\mathrm{Ind}(A) = k$. *Then*
(a) $r[\, AA^D - (AA^D)^*\,] = 2r[\, A^k,\ (A^k)^*\,] - 2r(A^k)$.
(b) $r[\, (AA^D)(AA^D)^* - (AA^D)^*(AA^D)\,] = 2r[\, A^k,\ (A^k)^*\,] - 2r(A^k)$.
(c) $r[\, AA^{\#} - (AA^{\#})^*\,] = r[\, (AA^{\#})(AA^{\#})^* - (AA^{\#})^*(AA^{\#})\,] = 2r[\, A,\ A^*\,] - 2r(A)$, *if* $\mathrm{Ind}(A) = 1$.
(d) $AA^D = (AA^D)^* \Leftrightarrow (AA^D)(AA^D)^* = (AA^D)^*(AA^D) \Leftrightarrow R(A^k) = R[(A^k)^*]$, *i.e., A^k is EP.*
(e) $AA^{\#} = (AA^{\#})^* \Leftrightarrow (AA^{\#})(AA^{\#})^* = (AA^{\#})^*(AA^{\#}) \Leftrightarrow R(A^*) = R(A)$, *i.e., A is EP.*

Proof. Note that both AA^D and $(AA^D)^*$ are idempotent. It follows from (3.1) that

$$\begin{aligned} r[\, AA^D - (AA^D)^*\,] &= r\begin{bmatrix} AA^D \\ (AA^D)^* \end{bmatrix} + r[\, AA^D,\ (AA^D)^*\,] - r(AA^D) - r[(AA^D)^*] \\ &= 2r[\, AA^D,\ (AA^D)^*\,] - 2r(A^D) \\ &= 2r[\, A^k,\ (A^k)^*\,] - 2r(A^k), \end{aligned}$$

as required for Part (a). Part (b) follows from Part (a) and Corollary 3.26(d). The results in Parts (c)—(e) follow immediately from Part (a). □

Theorem 13.10. *Let $A \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = k$. Then*

(a) $r(\,A^{\dagger} - A^{D}\,) = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} + r[\,A^k, \ A^*\,] - r(A^k) - r(A)$.

(b) $r(\,A^{\dagger} - A^{D}\,) = r(A^{\dagger}) - r(A^{D})$, *i.e.,* $A^D \leq_{rs} A^{\dagger} \Leftrightarrow R(A^k) \subseteq R(A^*)$ *and* $R[(A^k)^*] \subseteq r(A)$, *i.e.,* A *is power–EP.*

(c) $r(\,A^{\dagger} - A^{\#}\,) = 2r[\,A, \ A^*\,] - 2r(A)$, *If* $\mathrm{Ind}\,(A) = 1$.

(d)[16] $A^{\dagger} = A^{\#} \Leftrightarrow R(A^*) = R(A)$, *i.e.,* A *is EP.*

Proof. Since both $A^{\dagger}$ and A^D are outer inverses of A, it follows from (5.1) that

$$\begin{aligned} r(\,A^{\dagger} - A^{D}\,) &= r\begin{bmatrix} A^{\dagger} \\ A^D \end{bmatrix} + r[\,A^{\dagger}, \ A^D\,] - r(A^{\dagger}) - r(A^D) \\ &= r\begin{bmatrix} A^* \\ A^k \end{bmatrix} + r[\,A^*, \ A^k\,] - r(A) - r(A^k), \end{aligned}$$

as required for Part (a). The results in Part (b)—(d) follows immediately from Part (a). □

Theorem 13.11. *Let $A \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = k$. Then*

(a) $r(\,AA^{\dagger} - AA^{D}\,) = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} - r(A^k)$.

(b) $r(\,A^{\dagger}A - A^{D}A\,) = r[\,A^k, \ A^*\,] - r(A^k)$.

(c) $r[\,A^k(A^k)^{\dagger} - AA^D\,] = r\begin{bmatrix} A^k \\ (A^*)^k \end{bmatrix} - r(A^k)$.

(d) $r[\,(A^k)^{\dagger}A^k - A^DA\,] = r[\,A^k, \ (A^*)^k\,] - r(A^k)$.

In particular,

(e) $r(\,AA^{\dagger} - AA^{\#}\,) = r(\,A^{\dagger}A - A^{\#}A\,) = r[\,A, \ A^*\,] - r(A)$, *if* $\mathrm{Ind}\,(A) = 1$.

(f) $r(\,AA^{\dagger} - AA^{D}\,) = r(AA^{\dagger}) - r(AA^D)$, *i.e.,* $AA^D \leq_{rs} AA^{\dagger} \Leftrightarrow R[(A^k)^*] \subseteq R(A)$.

(g) $r(\,A^{\dagger}A - A^{D}A\,) = r(A^{\dagger}A) - r(A^DA)$, *i.e.,* $A^DA \leq_{rs} A^{\dagger}A \Leftrightarrow R(A^k) \subseteq R(A^*)$.

(h) $A^k(A^k)^{\dagger} = AA^D \Leftrightarrow (A^k)^{\dagger}A^k = A^DA \Leftrightarrow A^k$ *is EP.*

(i) $AA^{\dagger} = AA^{\#} \Leftrightarrow A^{\dagger}A = A^{\#}A \Leftrightarrow A$ *is EP.*

Proof. Note that both $AA^{\dagger}$ and AA^D are idempotent. Then it follows by (3.1) that

$$\begin{aligned} r(\,AA^{\dagger} - AA^{D}\,) &= r\begin{bmatrix} AA^{\dagger} \\ AA^D \end{bmatrix} + r[\,AA^{\dagger}, \ AA^D\,] - r(AA^{\dagger}) - r(AA^D) \\ &= r\begin{bmatrix} A^* \\ A^k \end{bmatrix} + r[\,A, \ A^k\,] - r(A) - r(A^k) = r\begin{bmatrix} A^* \\ A^k \end{bmatrix} - r(A^k), \end{aligned}$$

as required for Part (a). Similarly we can establish Parts (b)—(d). Part (e) is a special case of (b)—(d). Based on them we easily get Parts (f)—(i). □

Theorem 13.12. *Let $A \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = k$. Then*

(a) $r(\,A^{\dagger}A^D - A^DA^{\dagger}\,) = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} + r[\,A^k, \ A^*\,] - 2r(A) = r(\,A^{\dagger}A^k - A^kA^{\dagger}\,)$.

(b) $r(\,A^{\dagger}AA^DA - AA^DAA^{\dagger}\,) = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} + r[\,A^k, \ A^*\,] - 2r(A)$.

(c) $r(\,A^{\dagger}AA^D - A^DAA^{\dagger}\,) = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} + r[\,A^k, \ A^*\,] - 2r(A)$.

(d) $r(\,A^{\dagger}A^{\#} - A^{\#}A^{\dagger}\,) = 2r[\,A, \ A^*\,] - 2r(A)$, *if* $\mathrm{Ind}(A) = 1$.

(e) $r(\,A^{\dagger}AA^{\#} - A^{\#}AA^{\dagger}\,) = 2r[\,A, \ A^*\,] - 2r(A)$, *if* $\mathrm{Ind}(A) = 1$.

In particular,

(f) $A^{\dagger}$ *commutes with* $A^D \Leftrightarrow A^{\dagger}$ *commutes with* $AA^DA \Leftrightarrow A^{\dagger}$ *commutes with* $AA^D \Leftrightarrow A^{\dagger}$ *commutes with* $A^k \Leftrightarrow AA^{\dagger} = AA^D$ *and* $A^DA = A^{\dagger}A \Leftrightarrow R(A^k) \subseteq R(A^*)$ *and* $[(A^k)^*] \subseteq R(A)$, *i.e.,* A *is power-EP.*

(g) $A^\dagger$ *commutes with* $A^\#$ $\Leftrightarrow$ $A^\dagger$ *commutes with* $AA^\#$ $\Leftrightarrow$ A *is EP.*

Proof. Applying (2.2) to $A^\dagger A^D - A^D A^\dagger$ yields

$$\begin{aligned} r(\, A^\dagger A^D - A^D A^\dagger \,) &= r\begin{bmatrix} A^*AA^* & 0 & A^*A^D \\ 0 & -A^*AA^* & A^* \\ A^* & A^DA^* & 0 \end{bmatrix} - 2r(A) \\ &= r\begin{bmatrix} A^*AA^* & A^*A^DAA^* & A^*A^D \\ 0 & 0 & A^* \\ A^* & A^DA^* & 0 \end{bmatrix} - 2r(A) \\ &= r\begin{bmatrix} 0 & 0 & A^*A^D \\ 0 & 0 & A^* \\ A^* & A^DA^* & 0 \end{bmatrix} - 2r(A) \\ &= r\begin{bmatrix} A^D \\ A^* \end{bmatrix} + r[\, A^D, \ A^* \,] - 2r(A) = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} + r[\, A^k, \ A^* \,] - 2r(A). \end{aligned}$$

The second one in Part (a) is from Theorem 13.10(a) and (b). The two formulas in Part (b) and (c) are derived from (4.1) by noting that $AA^D = A^DA$ is idempotent. Parts (d)—(g) are direct consequences of Parts (a)—(c). □

Theorem 13.13. *Let* $A \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$. *Then*

(a) $r[\,(AA^\dagger)A^D - A^D(AA^\dagger)\,] = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} - r(A)$.

(b) $r[\,(A^\dagger A)A^D - A^D(A^\dagger A)\,] = r[\, A^k, \ A^* \,] - r(A)$.

(c) $r(\, A^\dagger A^D - A^DA^\dagger \,) = r[\,(AA^\dagger)A^D - A^D(AA^\dagger)\,] + r[\,(A^\dagger A)A^D - A^D(A^\dagger A)\,]$.

(d) $r[\,(AA^\dagger)A^\# - A^\#(AA^\dagger)\,] = r[\,(A^\dagger A)A^\# - A^\#(A^\dagger A)\,] = r[\, A, \ A^* \,] - r(A)$.

(e) A^D *commutes with* $AA^\dagger \Leftrightarrow R[(A^k)^*] \subseteq R(A)$.

(f) A^D *commutes with* $A^\dagger A \Leftrightarrow R(A^k) \subseteq R(A^*)$.

(g) $A^\dagger A^D = A^DA^\dagger$ $\Leftrightarrow$ A^D *commutes with* $A^\dagger A$ *and* A^D *commutes with* $A^\dagger A$ $\Leftrightarrow$ $R(A^k) \subseteq R(A^*)$ *and* $R(A^k) \subseteq R(A^*)$ $\Leftrightarrow$ A *is power-EP.*

(h) $A^\dagger A^\# = A^\# A^\dagger$ $\Leftrightarrow$ $A^\#$ *commutes with* $A^\dagger A$ $\Leftrightarrow$ $A^\#$ *commutes with* $A^\dagger A$ $\Leftrightarrow$ A *is EP.*

Proof. Note that both $AA^\dagger$ and $A^\dagger A$ are idempotent. Thus Parts (a) and (b) can easily be established through (4.1). Contrasting Parts (a) and (b) with Theorem 13.12(a) yields Part (c). Parts (d)—(g) are direct consequences of Parts (a)—(b). □

Theorem 13.14. *Let* $A \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$. *Then*

(a) $r(\, A^*A^D - A^DA^* \,) = r\begin{bmatrix} A^k(\, AA^* - A^*A\,)A^k & 0 & A^kA^* \\ 0 & 0 & A^k \\ A^*A^k & A^k & 0 \end{bmatrix} - 2r(A^k)$.

(b) $r(\, A^*A^\# - A^\#A^* \,) = r\begin{bmatrix} A(\, AA^* - A^*A\,)A & 0 & AA^* \\ 0 & 0 & A \\ A^*A & A & 0 \end{bmatrix} - 2r(A)$.

(c) $r(\, A^*A^D - A^DA^* \,) = r(\, A^{k+1}A^*A^k - A^kA^*A^{k+1} \,)$, *if* $R(A^*A^k) \subseteq R(A^k)$ *and* $R[A(A^k)^*] \subseteq R[(A^k)^*]$.

(d) $r(\, A^*A^D - A^DA^* \,) = r\begin{bmatrix} A^kA^* \\ A^k \end{bmatrix} + r[\, A^k, \ A^*A^k \,] - 2r(A^k)$, *if* $A^{k+1}A^*A^k = A^kA^*A^{k+1}$.

(e) $A^*A^D = A^DA^* \Leftrightarrow R(A^*A^k) \subseteq R(A^k)$, $R[A(A^k)^*] \subseteq R[(A^k)^*]$ *and* $A^{k+1}A^*A^k = A^kA^*A^{k+1}$.

(f) $r(\, A^*A^\# - A^\#A^* \,) = r(\, A^2A^*A - AA^*A^2 \,)$, *if* A *is EP.*

(g) $A^*A^\# = A^\#A^* \Leftrightarrow A^2A^*A = AA^*A^2$ *and* A *is EP* $\Leftrightarrow$ A *is both EP and star-dagger.*

Proof. Applying (2.3) to $A^*A^D - A^DA^*$ yields

$$\begin{aligned} r(\, A^*A^D - A^DA^* \,) &= r[\, A^*A^k(A^{2k+1})^\dagger A^k - A^k(A^{2k+1})^\dagger A^kA^* \,] \\ &= r\begin{bmatrix} -A^{2k+1} & 0 & A^k \\ 0 & A^{2k+1} & A^kA^* \\ A^*A^k & A^k & 0 \end{bmatrix} - 2r(A^{2k+1}) \end{aligned}$$

$$
\begin{aligned}
&= r\begin{bmatrix} -A^{2k+1} & 0 & A^k \\ -A^{k+1}A^*A^k & 0 & A^kA^* \\ A^*A^k & A^k & 0 \end{bmatrix} - 2r(A^k) \\
&= r\begin{bmatrix} 0 & 0 & A^k \\ A^kA^*A^{k+1} - A^{k+1}A^*A^k & 0 & A^kA^* \\ A^*A^k & A^k & 0 \end{bmatrix} - 2r(A^k) \\
&= r\begin{bmatrix} A^k(\,AA^* - A^*A\,)A^k & 0 & A^kA^* \\ 0 & 0 & A^k \\ A^*A^k & A^k & 0 \end{bmatrix} - 2r(A^k).
\end{aligned}
$$

as required for Part (a). Parts (b)—(d) are special cases of Part (a). Applying Lemma 1.2(f) to the rank equality in Part (a) yields Part (e). Parts (f) and (g) follow from Part (b). □

Similarly we can also establish the following four theorems, which proofs are omitted.

Theorem 13.15. *Let $A \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = k$. Then*

(a) $r(\,AA^*A^D - A^DA^*A\,) = r\begin{bmatrix} A^k(\,A^2A^* - A^*A^2\,)A^k & 0 & A^kA^*A \\ 0 & 0 & A^k \\ AA^*A^k & A^k & 0 \end{bmatrix} - 2r(A^k).$

(b) $r(\,A^kA^*A^D - A^DA^*A^k\,) = r(\,A^{2k+1}A^*A^k - A^kA^*A^{2k+1}\,).$

(c) $r(\,AA^*A^{\#} - A^{\#}A^*A\,) = r(\,A^3A^*A - AA^*A^3\,)$, *if* $\mathrm{Ind}(A) = 1$.

(d) $AA^*A^D = A^DA^*A \Leftrightarrow R(AA^*A^k) = R(A^k)$, $R[(A^kA^*A)^*] = R[(A^k)^*]$ *and* $A^{k+2}A^*A^k = A^kA^*A^{k+2}$.

(e) $A^kA^*A^D = A^DA^*A^k \Leftrightarrow A^{2k+1}A^*A^k = A^kA^*A^{2k+1}$.

(f)[64] *If A is star-dagger, then* $AA^*A^D = A^DA^*A$.

(g) $AA^*A^{\#} = A^{\#}A^*A \Leftrightarrow A^3A^*A = AA^*A^3$.

Theorem 13.16. *Let $A \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = k$. Then*

(a) $r(\,AA^DA^* - A^*A^DA\,) = r\begin{bmatrix} A^k \\ A^kA^* \end{bmatrix} + r[\,A^k,\ A^*A^k\,] - 2r(A^k) = r[\,A(A^k)^\dagger - (A^k)^\dagger A\,].$

(b) $r(\,AA^{\#}A^* - A^*A^{\#}A\,) = 2r[\,A,\ A^*\,] - 2r(A)$, *if* $\mathrm{Ind}(A) = 1$.

(c) $AA^DA^* = A^*A^DA \Leftrightarrow A(A^k)^\dagger = (A^k)^\dagger A \Leftrightarrow R(A^*A^k) = R(A^k)$ *and* $R[A(A^k)^*] = R[(A^k)^*]$.

(d) $AA^{\#}A^* = A^*A^{\#}A \Leftrightarrow A$ *is EP.*

Theorem 13.17. *Let $A \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = k$. Then*

(a) $r[\,AA^D(A^*)^k - (A^*)^kA^DA\,] = 2r[\,A^k,\ (A^k)^*\,] - 2r(A^k)$.

(b) $AA^D(A^*)^k = (A^*)^kA^DA \Leftrightarrow A^k$ *is EP.*

Theorem 13.18. *Let $A \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = 1$ and λ is a nonzero complex number. Then*

(a) $r[\,AA^{\#}(\,AA^* + \lambda A^*A\,) - (\,AA^* + \lambda A^*A\,)A^{\#}A\,] = 2r[\,A,\ A^*\,] - 2r(A)$.

(b) $AA^{\#}$ *commutes with* $AA^* + \lambda A^*A \Leftrightarrow A$ *is EP.*

Theorem 13.19. *Let $A \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = k$. Then*

(a) $r[\,(AA^D)^*A^\dagger - A^\dagger(AA^D)^*\,] = r\begin{bmatrix} A^kA^*A \\ A^k \end{bmatrix} + r[\,AA^*A^k,\ A^k\,] - 2r(A^k)$.

(b) $(AA^D)^*A^\dagger = A^\dagger(AA^D)^* \Leftrightarrow R(AA^*A^k) = R(A^k)$ *and* $R[(A^kA^*A)^*] = R[(A^k)^*]$.

(c) $(AA^{\#})^*A^\dagger = A^\dagger(AA^{\#})^*$, *if* $\mathrm{Ind}(A) = 1$.

Proof. Apply (4.1) to $(AA^D)^*A^\dagger - A^\dagger(AA^D)^*$ to yield

$$
\begin{aligned}
r[\,(AA^D)^*A^\dagger - A^\dagger(AA^D)^*\,] &= r\begin{bmatrix} (AA^D)^*A^\dagger \\ (AA^D)^* \end{bmatrix} + r[\,A^\dagger(AA^D)^*,\ (AA^D)^*\,] - 2r(AA^D) \\
&= r\begin{bmatrix} (A^k)^*A^\dagger \\ (A^k)^* \end{bmatrix} + r[\,A^\dagger(A^k)^*,\ (A^k)^*\,] - 2r(A^k).
\end{aligned}
$$

Next applying (2.1), we find that

$$
r\begin{bmatrix} (A^k)^*A^\dagger \\ (A^k)^* \end{bmatrix} = r[\,AA^*A^k,\ A^k\,] \ \text{ and } \ r[\,A^\dagger(A^k)^*,\ (A^k)^*\,] = r\begin{bmatrix} A^kA^*A \\ A^k \end{bmatrix}.
$$

Thus we get Part (a). □

Theorem 13.20. *Let $A \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = k$. Then*

(a) $r[\, A - (A^D)^\dagger \,] = r\begin{bmatrix} A & A^k & (A^k)^* \\ A^k & 0 & 0 \\ (A^k)^* & 0 & 0 \end{bmatrix} - 2r(A^k).$

(b) $r[\, A - (A^D)^\dagger \,] = r(A) - r(A^k)$, *if A^k is EP.*

(c) $(A^{\#})^\dagger = A \Leftrightarrow A$ *is EP.*

Proof. According to Cline's identity $(A^D)^\dagger = (A^k)^\dagger A^{2k+1}(A^k)^\dagger$(see [16] and [32]), we find by (2.8) that

$$\begin{aligned} r[\, A - (A^D)^\dagger \,] &= r[\, A - (A^k)^\dagger A^{2k+1}(A^k)^\dagger \,] \\ &= r\begin{bmatrix} (A^k)^*A^{2k+1}(A^k)^* & (A^k)^*A^k(A^k)^* & 0 \\ (A^k)^*A^k(A^k)^* & 0 & (A^k)^* \\ 0 & (A^k)^* & -A \end{bmatrix} - 2r(A^k) \\ &= r\begin{bmatrix} A^{2k+1} & A^k(A^k)^* & 0 \\ (A^k)^*A^k & 0 & (A^k)^* \\ 0 & (A^k)^* & -A \end{bmatrix} - 2r(A^k) \\ &= r\begin{bmatrix} A^{2k+1} & 0 & A^{k+1} \\ (A^k)^*A^k & 0 & (A^k)^* \\ 0 & (A^k)^* & -A \end{bmatrix} - 2r(A^k) \\ &= r\begin{bmatrix} 0 & 0 & A^{k+1} \\ 0 & 0 & (A^k)^* \\ A^{k+1} & (A^k)^* & -A \end{bmatrix} - 2r(A^k) \\ &= r\begin{bmatrix} A & A^k & (A^k)^* \\ A^k & 0 & 0 \\ (A^k)^* & 0 & 0 \end{bmatrix} - 2r(A^k). \end{aligned}$$

as required for Part (a). The results in Part (b) and (c) follows immediately from Part (a). □

Theorem 13.21. *Let $A \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = k$. Then*

(a) $r[\, AA^DA - (A^D)^\dagger \,] = 2r[\, A^k,\ (A^k)^* \,] - 2r(A^k).$

(b) $(A^D)^\dagger = AA^DA \Leftrightarrow A^k$ *is EP.*

Proof. It is easy to verify that both AA^DA and $(A^D)^\dagger$ are outer inverses of A^D. In that case it follows from (5.1) that

$$\begin{aligned} r[\, AA^DA - (A^D)^\dagger \,] &= r\begin{bmatrix} AA^DA \\ (A^D)^\dagger \end{bmatrix} + r[\, AA^DA,\ (A^D)^\dagger \,] - r(AA^DA) - r[(A^D)^\dagger] \\ &= r\begin{bmatrix} A^D \\ (A^D)^* \end{bmatrix} + r[\, A^D,\ (A^D)^* \,] - 2r(A^k) \\ &= r\begin{bmatrix} A^k \\ (A^k)^* \end{bmatrix} + r[\, A^k,\ (A^k)^* \,] - 2r(A^k), \end{aligned}$$

as required for Part (a). The result in Part (b) follows immediately from Part (a). □

Theorem 13.22. *Let $A,\ B \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(A) = k$ and $\mathrm{Ind}(B) = l$. Then*

(a) $r(\, AA^D - BB^D \,) = r\begin{bmatrix} A^k \\ B^l \end{bmatrix} + r[\ A^k,\ B^l \,] - r(A^k) - r(B^l).$

(b) $r(\, AA^{\#} - BB^{\#} \,) = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\ A,\ B \,] - r(A) - r(B)$, *If* $\mathrm{Ind}(A) = \mathrm{Ind}(B) = l$.

(c) $AA^D = BB^D \Leftrightarrow R(A^k) = R(B^l)$ *and* $R[(A^k)^*] = R[(B^l)^*]$.

(d) $r(\, AA^D - BB^D \,)$ *is nonsingular* $\Leftrightarrow r\begin{bmatrix} A^k \\ B^l \end{bmatrix} = r[\ A^k,\ B^l \,] = r(A^k) + r(B^l) = m \Leftrightarrow R(A^k) \oplus R(B^l) = R[(A^k)^*] \oplus R[(B^l)^*] = \mathcal{C}^m$.

(e) *In particular, if* $\operatorname{Ind}\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = 1$. *Then*

$$r\left(\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}^{\#} - \begin{bmatrix} AA^{\#} & 0 \\ 0 & DD^{\#} \end{bmatrix} \right) = r[\, A, \ B\,] + r\begin{bmatrix} B \\ D \end{bmatrix} - r\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}.$$

Proof. Note that both AA^D and BB^D are idempotent. Then it follows from (3.1) that

$$\begin{aligned} r(\, AA^D - BB^D\,) &= r\begin{bmatrix} AA^D \\ BB^D \end{bmatrix} + r[\, AA^D, \ BB^D\,] - r(AA^D) - r(BB^D) \\ &= r\begin{bmatrix} A^D \\ B^D \end{bmatrix} + r[\, A^D, \ B^D\,] - r(A^D) - r(B^D) \\ &= r\begin{bmatrix} A^k \\ B^l \end{bmatrix} + r[\, A^k, \ B^l\,] - r(A^k) - r(B^l), \end{aligned}$$

as required for Part (a). The results in Parts (b)—(e) follow immediately from Part (a). □

Theorem 13.23. *Let* $A, B \in \mathcal{C}^{m\times m}$ *with* $\operatorname{Ind}(A) = k$. *Then*

(a) $r(\, AA^DB - BA^DA\,) = r\begin{bmatrix} A^k \\ A^kB \end{bmatrix} + r[\, A^k, \ BA^k\,] - 2r(A^k)$.

(b) $r(\, A^DAA^\dagger - A^\dagger AA^D\,) = r\begin{bmatrix} A^k \\ A^* \end{bmatrix} + r[\, A^k, \ A^*\,] - 2r(A) = r(\, A^DA^\dagger - A^\dagger A^D\,)$.

In particular,

(c) $AA^DB = BA^DA \Leftrightarrow R(BA^k) = R(A^k)$ *and* $R[(A^kB)^*] = R[(A^k)^*]$.

(d) $A^DAA^\dagger = A^\dagger AA^D \Leftrightarrow A^DA^\dagger = A^\dagger A^D \Leftrightarrow R(A^k) \subseteq R(A^*)$ *and* $R[(A^k)^*] \subseteq R(A)$.

Proof. Note that $AA^D = A^DA$ is idempotent. It follows by (4.1) that

$$\begin{aligned} r(\, AA^DB - BA^DA\,) &= r\begin{bmatrix} AA^DB \\ A^DA \end{bmatrix} + r[\, BA^DA, \ AA^D\,] - r(AA^D) - r(A^DA) \\ &= r\begin{bmatrix} A^DB \\ A^D \end{bmatrix} + r[\, BA^D, \ A^D\,] - 2r(A^D) \\ &= r\begin{bmatrix} A^kB \\ A^k \end{bmatrix} + r[\, BA^k, \ A^k\,] - 2r(A^k). \end{aligned}$$

Thus we have Parts (a). Replacing B by $A^\dagger$ in Part (a) and simplifying it yields the first equality in Part (b). The second equality in Part (b) follows from Theorem 13.22(a). □

Theorem 13.24. *Let* $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times m}$ *with* $\operatorname{Ind}(B) = k$ *and* $C \in \mathcal{C}^{n\times n}$ *with* $\operatorname{Ind}(C) = l$. *Then*

(a) $r(\, BB^DA - AC^DC\,) = r\begin{bmatrix} B^kA \\ C^l \end{bmatrix} + r[\, AC^l, \ B^k\,] - r(B^k) - r(C^l)$.

(b) $BB^DA = AC^DC \Leftrightarrow R(AC^l) \subseteq R(B^k)$ *and* $R[(B^kA)^*] \subseteq R[(C^l)^*]$.

Proof. Note that both BB^D and C^DC are idempotent. Then it follows from (4.1) that

$$\begin{aligned} r(\, BB^DA - AC^DC\,) &= r\begin{bmatrix} BB^DA \\ C^DC \end{bmatrix} + r[\, AC^DC, \ BB^D\,] - r(BB^D) - r(C^DC) \\ &= r\begin{bmatrix} B^DA \\ C^D \end{bmatrix} + r[\, AC^D, \ B^D\,] - r(B^D) - r(C^D) \\ &= r\begin{bmatrix} B^kA \\ C^l \end{bmatrix} + r[\, AC^l, \ B^k\,] - r(B^k) - r(C^l), \end{aligned}$$

as required for Part (a). □

Theorem 13.25. *Let* $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times m}$ *with* $\operatorname{Ind}(B) = k$ *and* $C \in \mathcal{C}^{n\times n}$ *with* $\operatorname{Ind}(C) = l$. *Then*

(a) $r[\,A,\ B^k\,] = r(B^k) + r(\,A - BB^DA\,)$.

(b) $r\begin{bmatrix} A \\ C^l \end{bmatrix} = r(C^l) + r(\,A - AC^DC\,)$.

(c) $r\begin{bmatrix} A & B^k \\ C^l & 0 \end{bmatrix} = r(B^k) + r(C^l) + r[\,(\,I_m - BB^D\,)A(\,I_n - C^DC\,)\,]$.

Proof. Applying (1.7) to $A - BB^DA$ yields

$$\begin{aligned} r(\,A - BB^DA\,) &= r[\,A - B^{k+1}(B^{2k+1})^{\dagger}B^kA\,] \\ &= r\begin{bmatrix} B^{2k+1} & B^kA \\ B^{k+1} & A \end{bmatrix} - r(B^{2k+1}) \\ &= r\begin{bmatrix} 0 & 0 \\ B^{k+1} & A \end{bmatrix} - r(B^k) = r[\,A,\ B^k\,] - r(B^k), \end{aligned}$$

as required for Part (a). Similarly we can show Parts (b) and (c). □

Theorem 13.26. *Let* $A,\ B \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$ *and* $\mathrm{Ind}(B) = l$. *Then*

(a) $r(\,AB - ABB^DA^DAB\,) = r\begin{bmatrix} A^{2k} & A^kB^l \\ B^lA^k & B^{2l} \end{bmatrix} + r(AB) - r(A^k) - r(B^l)$.

(b) $r(\,AB - ABB^{\#}A^{\#}AB\,) = r\begin{bmatrix} A^2 & AB \\ BA & B^2 \end{bmatrix} + r(AB) - r(A) - r(B)$, *if* $\mathrm{Ind}(A) = \mathrm{Ind}(B) = 1$.

(c) $B^DA^D \in \{(AB)^-\} \Leftrightarrow r\begin{bmatrix} A^{2k} & A^kB^l \\ B^lA^k & B^{2l} \end{bmatrix} = r(A^k) + r(B^l) - r(AB)$.

(d) $B^{\#}A^{\#} \in \{(AB)^-\} \Leftrightarrow r\begin{bmatrix} A^2 & AB \\ BA & B^2 \end{bmatrix} = r(A) + r(B) - r(AB)$.

Proof. It follows by (2.9) that

$$\begin{aligned} r(\,AB - ABB^DA^DAB\,) &= r[\,AB - AB^{k+1}(B^{2k+1})^{\dagger}B^lA^k(A^{2k+1})^{\dagger}A^{k+1}B\,] \\ &= r\begin{bmatrix} B^lA^k & B^{2l+1} & 0 \\ A^{2k+1} & 0 & A^{k+1}B \\ 0 & AB^{l+1} & -AB \end{bmatrix} - r(A^{2k+1}) - r(B^{2l+1}) \\ &= r\begin{bmatrix} B^lA^k & B^{2l+1} & 0 \\ A^{2k+1} & A^{k+1}B^{l+1} & 0 \\ 0 & 0 & -AB \end{bmatrix} - r(A^k) - r(B^l) \\ &= r\begin{bmatrix} B^lA^k & B^{2l+l} \\ A^{2k+1} & A^kB^{l+l} \end{bmatrix} + r(AB) - r(A^k) - r(B^l) \\ &= r\begin{bmatrix} B^lA^k & B^{2l} \\ A^{2k} & A^kB^l \end{bmatrix} + r(AB) - r(A^k) - r(B^l). \end{aligned}$$

Thus we have Parts (a). □

Theorem 13.27. *Let* $A,\ B \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$ *and* $\mathrm{Ind}(B) = l$. *Then*

(a) $r(\,AA^DB^DB - BB^DA^DA\,) = r\begin{bmatrix} A^k \\ B^l \end{bmatrix} + r[\,A^k,\ B^l\,] + r(A^kB^l) + r(B^lA^k) - 2r(A^k) - 2r(B^l)$.

(b) $r(\,AA^{\#}B^{\#}B - BB^{\#}A^{\#}A\,) = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\,A,\ B\,] + r(AB) + r(BA) - 2r(A) - 2r(B)$, *if* $\mathrm{Ind}(A) = \mathrm{Ind}(B) = 1$.

(c) $AA^DB^DB = BB^DA^DA \Leftrightarrow r\begin{bmatrix} A^k \\ B^l \end{bmatrix} = r(A^k) + r(B^l) - r(A^kB^l)$ *and* $r[\,A^k,\ B^l\,] = r(A^k) + r(B^l) - r(B^lA^k)$.

(d) $AA^{\#}B^{\#}B = BB^{\#}A^{\#}A \Leftrightarrow r\begin{bmatrix} A \\ B \end{bmatrix} = r(A) + r(B) - r(AB)$ *and* $r[\,A,\ B\,] = r(A) + r(B) - r(BA)$.

Proof. Note that both $AA^D = A^DA$ and $BB^D = B^DB$ are idempotent. Then it follows by (3.26) that

$$r(\,AA^DB^DB - BB^DA^DA\,)$$

$$
\begin{aligned}
&= r\begin{bmatrix} AA^D \\ BB^D \end{bmatrix} + r[\, A^DA, \ B^DB\,] + r(AA^DB^DB) + r(BB^DA^DA) - 2r(AA^D) - 2r(BB^D) \\
&= r\begin{bmatrix} A^k \\ B^l \end{bmatrix} + r[\, A^k, \ B^l\,] + r(A^kB^l) + r(B^lA^k) - 2r(A^k) - 2r(B^l),
\end{aligned}
$$

as required for Part (a). □

Theorem 13.28. *Let* $A, B \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(\, A+B\,) = k$ *and denote* $N = A+B$. *Then*

(a) $r(AN^DB) = r(AN^k) + r(N^kB) - r(N^k)$.

(b) $r(\, AN^DB\,) = r(A) + r(B) - r(N^k)$, *if* $R(B) \subseteq R(N^k)$ *and* $R(A^*) \subseteq R[(N^k)^*]$.

(c) $r(\, AN^DB - BN^DA\,) = r\begin{bmatrix} N^k \\ N^kB \end{bmatrix} + r[\, N^k, \ BN^k\,] - 2r(N^k)$.

(d) $AN^DB = BN^DA \Leftrightarrow R(BN^k) \subseteq R(N^k)$ *and* $R[(N^kB)^*] \subseteq R[(N^k)^*]$.

Proof. It follows by (1.7), that

$$
\begin{aligned}
r(AN^DB) &= r[\, AN^k(N^{2k+1})^{\dagger}N^kB\,] \\
&= r\begin{bmatrix} N^{2k+1} & N^kB \\ AN^k & 0 \end{bmatrix} - r(N^{2k+1}) \\
&= r\begin{bmatrix} 0 & N^kB \\ AN^k & 0 \end{bmatrix} - r(N^k) = r(AN^k) + r(N^kB) - r(N^k),
\end{aligned}
$$

which is the first equality in Part (a). The second equality in Part (a) follows from $r(AN^k) = r(AN^D)$, $r(N^kB) = r(N^DB)$ and $r(N^D) = r(N^k)$. Under $R(B) \subseteq R(N^k)$ and $R(A^*) \subseteq R[(N^k)^*]$, it follows that $r(AN^k) = r(A)$ and $r(N^kB) = r(B)$. Thus Part (a) becomes Part (b). Next applying (2.3) to $AN^DB - BN^DA$ yields

$$
\begin{aligned}
&r(\, AN^DB - BN^DA\,) \\
&= r[\, AN^k(N^{2k+1})^{\dagger}N^kB - BN^k(N^{2k+1})^{\dagger}N^kA\,] \\
&= r\begin{bmatrix} -N^{2k+1} & 0 & N^kB \\ 0 & N^{2k+1} & N^kA \\ AN^k & BN^k & 0 \end{bmatrix} - 2r(N^{2k+1}) \\
&= r\begin{bmatrix} -N^kAN^k & -N^kBN^k & N^kB \\ N^kAN^k & N^kBN^k & N^kA \\ AN^k & BN^k & 0 \end{bmatrix} - 2r(N^k) \\
&= r\begin{bmatrix} 0 & 0 & N^kB \\ 0 & 0 & N^kA \\ AN^k & BN^k & 0 \end{bmatrix} - 2r(N^k) \\
&= r\begin{bmatrix} N^kA \\ N^kB \end{bmatrix} + r[\, AN^k, \ BN^k\,] - 2r(N^k) = r\begin{bmatrix} N^k \\ N^kB \end{bmatrix} + r[\, N^k, \ BN^k\,] - 2r(N^k).
\end{aligned}
$$

Thus we have Parts (c) and (d). □

Theorem 13.29. *Let* $A, B \in \mathcal{C}^{m\times m}$ *be given, and let* $N = A+B$ *with* $\mathrm{Ind}(\, A+B\,) = k$. *Then*

(a) $r\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} - \begin{bmatrix} A \\ B \end{bmatrix}(\, A+B\,)^D[\, A, \ B\,] \right) = r(A) + r(B) - r(N^k)$.

(b) $\begin{bmatrix} A \\ B \end{bmatrix}(\, A+B\,)^D[\, A, \ B\,] = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Leftrightarrow \mathrm{Ind}(\, A+B\,) \leq 1$ *and* $r(\, A+B\,) = r(A) + r(B)$.

Proof. It follows by (1.7) that

$$
\begin{aligned}
&r\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} - \begin{bmatrix} A \\ B \end{bmatrix} N^D[\, A, \ B\,] \right) \\
&= r\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} - \begin{bmatrix} A \\ B \end{bmatrix} N^k(N^{2k+1})^{\dagger}N^k[\, A, \ B\,] \right)
\end{aligned}
$$

$$= r\begin{bmatrix} N^{2k+1} & N^kA & N^kB \\ AN^k & A & 0 \\ BN^k & 0 & B \end{bmatrix} - r(N^k)$$

$$= r\begin{bmatrix} 0 & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & B \end{bmatrix} - r(N^k) = r(A) + r(B) - r(N^k),$$

which is exactly Part (a). Note that $r(N^k) \leq r(N) = r(\, A+B\,) \leq r(A)+r(B)$. Thus $r(N^k) = r(A)+r(B)$ is equivalent to $\mathrm{Ind}(N) \leq 1$ and $r(N) = r(A) + r(B)$. $\square$

In general we have the following.

Theorem 13.30. *Let $A_1,\, A_2,\, \cdots,\, A_k \in \mathcal{C}^{m\times m}$ with $\mathrm{Ind}(N) = k$, where $N = A_1 + A_2 + \cdots + A_k$, and let $A = \mathrm{diag}(\, A_1,\, A_2,\, \cdots,\, A_k\,)$. Then*

(a) $r\left(A - \begin{bmatrix} A_1 \\ \vdots \\ A_k \end{bmatrix} N^D[\, A_1,\, \cdots,\, A_k\,] \right) = r(A_1) + \cdots + r(A_k) - r(N^k).$

(b) $\begin{bmatrix} A_1 \\ \vdots \\ A_k \end{bmatrix} N^D[\, A_1,\, \cdots,\, A_k\,] = A \Leftrightarrow \mathrm{Ind}(N) \leq 1$ *and* $r(N) = r(A_1) + \cdots + r(A_k)$.

Theorem 13.31. *Let $A_1,\, A_2,\, \cdots,\ ,A_k \in \mathcal{C}^{m\times m}$. Then the Drazin inverse of their sum satisfies the following equality*

$$(\, A_1 + A_2 + \cdots + A_k\,)^D = \frac{1}{k}[\, I_m,\, I_m,\, \cdots,\, I_m\,] \begin{bmatrix} A_1 & A_2 & \cdots & A_k \\ A_k & A_1 & \cdots & A_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_2 & A_3 & \cdots & A_1 \end{bmatrix}^D \begin{bmatrix} I_m \\ I_m \\ \vdots \\ I_m \end{bmatrix}. \tag{13.1}$$

Proof. Since the given matrices are square, (11.7) can be written as

$$U_m^* A U_m = \mathrm{diag}(\, J_1,\, J_2,\, \cdots,\, J_k\,).$$

In that case, it is easy to verify that

$$(U_m^* A U_m)^D = U_m^* A^D U_m,$$

and

$$[\,\mathrm{diag}(\, J_1,\, J_2,\, \cdots,\, J_k\,)\,]^D = \mathrm{diag}(\, J_1^D,\, J_2^D,\, \cdots,\, J_k^D\,)$$

Thus we have

$$J_1^D = [\, I_m,\, 0,\, \cdots,\, 0\,]U_m^* A^D U_m[\, I_m,\ 0,\ \cdots,\ 0\,]^T = \frac{1}{k}[\, I_m,\, I_m,\, \cdots,\, I_m\,]A^D[\, I_m,\, I_m,\, \cdots,\, I_m\,]^T,$$

which is (13.1). $\square$

Theorem 13.32. *Let $A + iB \in \mathcal{C}^{m\times m}$, where A and B are real. Then the Drazin inverse of $A + iB$ satisfies the identity*

$$(\, A + iB\,)^D = \frac{1}{2}[\, I_m,\ iI_m\,] \begin{bmatrix} A & -B \\ B & A \end{bmatrix}^D \begin{bmatrix} I_m \\ -iI_m \end{bmatrix}. \tag{13.2}$$

Proof. Observe that

$$\begin{bmatrix} A & iB \\ iB & A \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & iI_m \end{bmatrix} \begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & iI_m \end{bmatrix}^{-1}.$$

Thus

$$\begin{bmatrix} A & iB \\ iB & A \end{bmatrix}^D = \begin{bmatrix} I_m & 0 \\ 0 & iI_m \end{bmatrix} \begin{bmatrix} A & -B \\ B & A \end{bmatrix}^D \begin{bmatrix} I_m & 0 \\ 0 & iI_m \end{bmatrix}^{-1}. \tag{13.3}$$

In that case, applying (13.1) and then (13.3) to $A + iB$ yields (13.2). □

The identities in (11.31)—(11.33) can also be extended to the Drazin inverse of a real quaternion matrix.

Theorem 13.33. *Let $A_0 + iA_1 + iA_2 + kA_3$ be an $m \times m$ real quaternion matrix. Then its Drazin inverse satisfies*

$$(A_0 + iA_1 + iA_2 + kA_3)^D = \frac{1}{2}[\, I_m, \ jI_m \,] \begin{bmatrix} A_0 + iA_1 & -(A_2 + iA_3) \\ A_2 - iA_3 & A_0 - iA_1 \end{bmatrix}^D \begin{bmatrix} I_m \\ -jI_m \end{bmatrix}, \tag{13.4}$$

and

$$(A_0 + iA_1 + iA_2 + kA_3)^D = \frac{1}{4}[\, I_m, \ iI_m, \ jI_m, \ kI_m \,] \begin{bmatrix} A_0 & -A_1 & -A_2 & -A_3 \\ A_1 & A_0 & A_3 & -A_2 \\ A_2 & -A_3 & A_0 & A_1 \\ A_3 & A_2 & -A_1 & A_0 \end{bmatrix}^D \begin{bmatrix} I_m \\ -iI_m \\ -jI_m \\ -kI_m \end{bmatrix}. \tag{13.5}$$

Moreover denote $(A_0 + iA_1 + iA_2 + kA_3)^D = G_0 + iG_1 + iG_2 + kG_3$. Then

$$\begin{bmatrix} A_0 & -A_1 & -A_2 & -A_3 \\ A_1 & A_0 & A_3 & -A_2 \\ A_2 & -A_3 & A_0 & A_1 \\ A_3 & A_2 & -A_1 & A_0 \end{bmatrix}^D = \begin{bmatrix} G_0 & -G_1 & -G_2 & -G_3 \\ G_1 & G_0 & G_3 & -G_2 \\ G_2 & -G_3 & G_0 & G_1 \\ G_3 & G_2 & -G_1 & G_0 \end{bmatrix}^D. \tag{13.6}$$

As is well known that Drazin inverses of block matrices are quite difficult to determine in general. However, if a block matrix has some special pattern, its Drazin inverse can still be presented. Motivated by the expressions (9.78) and (8.82)—(9.85), we can find the following.

Let

$$M = \begin{bmatrix} A & B & \cdots & B \\ B & A & \cdots & B \\ \vdots & \vdots & \ddots & \vdots \\ B & B & \cdots & A \end{bmatrix}_{k \times k}, \tag{13.7}$$

where both A and B are $m \times m$ matrices. Then

$$M^D = \begin{bmatrix} S & T & \cdots & T \\ T & S & \cdots & T \\ \vdots & \vdots & \ddots & \vdots \\ T & T & \cdots & S \end{bmatrix}_{k \times k}, \tag{13.8}$$

where

$$S = \frac{1}{k}[\, A + (k-1)B \,]^D + \frac{k-1}{k}(\, A - B \,)^D, \qquad T = \frac{1}{k}[\, A + (k-1)B \,]^D - \frac{1}{k}(\, A - B \,)^D. \tag{13.9}$$

In fact we see from (9.77) that $M = P_m N P_m^{-1}$, when $m = n$. In that case, $M^D = P_m N^D P_m^{-1}$ holds. Written in an explicit form, it is (13.7). The expression (9.78) illustrates that M^D also has the same pattern as M.

Chapter 14

Rank equalities for submatrices in Drazin inverses

Let

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{14.1}$$

be a square block matrix over $\mathcal{C}$, where $A \in \mathcal{C}^{m\times m}$ and $D \in \mathcal{C}^{n\times n}$,

$$V_1 = \begin{bmatrix} A \\ C \end{bmatrix}, \quad V_2 = \begin{bmatrix} B \\ D \end{bmatrix}, \quad W_1 = [\, A, \ B \,], \quad W_2 = [\, C, \ D \,], \tag{14.2}$$

and partition the Drazin inverse of M as

$$M^D = \begin{bmatrix} G_1 & G_2 \\ G_3 & G_4 \end{bmatrix}, \tag{14.3}$$

where $G_1 \in \mathcal{C}^{m\times m}$. It is, in general, quite difficult to give the expression of G_1—G_4. In this chapter we consider a simpler problem—the ranks of the submatrices G_1—G_4 in (14.3).

Theorem 14.1. *Let M and M^D be given by* (14.1) *and* (14.3) *with* $\mathrm{Ind}(M) \geq 1$. *Then the ranks of G_1—G_4 in* (14.3) *can be determined by the following formulas*

$$r(G_1) = r\begin{bmatrix} M^kJ_1M^k & M^{k-1}V_1 \\ W_1M^{k-1} & 0 \end{bmatrix} - r(M^k), \tag{14.4}$$

$$r(G_2) = r\begin{bmatrix} M^kJ_2M^k & M^{k-1}V_2 \\ W_1M^{k-1} & 0 \end{bmatrix} - r(M^k), \tag{14.5}$$

$$r(G_3) = r\begin{bmatrix} M^kJ_3M^k & M^{k-1}V_1 \\ W_2M^{k-1} & 0 \end{bmatrix} - r(M^k), \tag{14.6}$$

$$r(G_4) = r\begin{bmatrix} M^kJ_4M^k & M^{k-1}V_2 \\ W_2M^{k-1} & 0 \end{bmatrix} - r(M^k), \tag{14.7}$$

where $V_1, \ V_2, \ W_1$ and W_2 are defined in (14.2), *and*

$$J_1 = \begin{bmatrix} -A & 0 \\ 0 & D \end{bmatrix}, \quad J_2 = \begin{bmatrix} 0 & B \\ -C & 0 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 0 & -B \\ C & 0 \end{bmatrix}, \quad J_4 = \begin{bmatrix} A & 0 \\ 0 & -D \end{bmatrix}. \tag{14.8}$$

Proof. We only show (14.4). In fact G_1 in (14.3) can be written as

$$G_1 = [\, I_m, \ 0 \,]M^D \begin{bmatrix} I_m \\ 0 \end{bmatrix} = P_1M^DQ_1 = P_1M^k(M^{2k+1})^\dagger M^kQ_1,$$

where $P_1 = [\, I_m, \ 0\,]$ and $Q_1 = \begin{bmatrix} I_m \\ 0 \end{bmatrix}$. Then it follows by Eq.(1.6) and block elementary operations that

$$\begin{aligned} r(G_1) &= r\begin{bmatrix} M^{2k+1} & M^kQ_1 \\ P_1M^k & 0 \end{bmatrix} - r(M^{2k+1}) \\ &= r\begin{bmatrix} M^{2k+1} - M^kQ_1P_1MM^k - M^kMQ_1P_1M^k & M^kQ_1 \\ P_1M^k & 0 \end{bmatrix} - r(M^k) \\ &= r\begin{bmatrix} M^k(\, M - Q_1P_1M - MQ_1P_1\,)M^k & M^kQ_1 \\ P_1M^k & 0 \end{bmatrix} - r(M^k) \\ &= r\begin{bmatrix} M^kJ_1M^k & M^{k-1}V_1 \\ W_1M^{k-1} & 0 \end{bmatrix} - r(M^k), \end{aligned}$$

which is exactly the equality (14.4). □

The further simplification of (14.4)—(14.7) is quite difficult, because the powers of M occur in them. However if M in (14.1) satisfies some additional conditions, the four rank equalities in (14.4)—(14.7) can reduce to simpler forms. We next present some of them. The first one is related to the well-known result on the Drazin inverse of an upper triangular block matrix (see Campbell and Meyer [21]).

$$\begin{bmatrix} A & B \\ 0 & N \end{bmatrix}^D = \begin{bmatrix} A^D & X \\ 0 & N^D \end{bmatrix}, \tag{14.9}$$

where

$$X = (A^D)^2\left[\sum_{i=0}^{l-1}(A^D)^iBN^i\right](\, I_n - N^DN\,) + (\, I_m - AA^D\,)\left[\sum_{i=0}^{k-1}A^iB(N^D)^i\right](N^D)^2 - A^DBN^D, \tag{14.10}$$

and $\mathrm{Ind}(A) = k, \ \mathrm{Ind}(N) = l$.

Theorem 14.2. *The rank of the submatrix X in* (14.9) *is*

$$r(X) = r\begin{bmatrix} A^k & P_t(B) & 0 \\ 0 & A^tBN^t & P_t(B) \\ 0 & 0 & N^l \end{bmatrix} - r\begin{bmatrix} A^k & P_t(B) \\ 0 & N^l \end{bmatrix}, \tag{14.11}$$

where $t = \mathrm{Ind}\begin{bmatrix} A & B \\ 0 & N \end{bmatrix}$, $P_t(B) = \sum_{i=0}^{t-1} A^{t-i-1}BN^i$. *In particular if $A^kBN^l = 0$, then*

$$r(X) = r[\, A^k, \ P_t(B)\,] + r\begin{bmatrix} P_t(B) \\ N^l \end{bmatrix} - r\begin{bmatrix} A^k & P_t(B) \\ 0 & N^l \end{bmatrix}. \tag{14.12}$$

In particular if $R[P_t(B)] \subseteq R(A^k)$ and $R[(P_t(B))^*] \subseteq R[(N^l)^*]$, then $r(X) = r(A^kBC^l)$.

Proof. It is easy to verify that

$$M^t = \begin{bmatrix} A & B \\ 0 & N \end{bmatrix}^t = \begin{bmatrix} A^t & P_t(B) \\ 0 & N^t \end{bmatrix}, \quad \text{and} \quad P_{2t+1}(B) = A^{t+1}P_t(B) + P_t(B)N^{t+1} + A^tBN^t.$$

Then applying (1.7) to $X = [\, I_m, \ 0\,]\begin{bmatrix} A^D & X \\ 0 & N^D \end{bmatrix}\begin{bmatrix} 0 \\ I_n \end{bmatrix} = P_1M^t(M^{2t+1})^\dagger M^tQ_2$, we find that

$$\begin{aligned} r(X) &= r\begin{bmatrix} M^{2t+1} & M^tQ_2 \\ P_1M^t & 0 \end{bmatrix} - r(M^{2k+1}) \\ &= r\begin{bmatrix} A^{2t+1} & P_{2t+1}(B) & P_t(B) \\ 0 & N^{2t+1} & N^t \\ A^t & P_t(B) & 0 \end{bmatrix} - r(M^k) \\ &= r\begin{bmatrix} 0 & A^tBN^t & P_t(B) \\ 0 & 0 & N^t \\ A^t & P_t(B) & 0 \end{bmatrix} - r(M^k) \end{aligned}$$

$$= \; r\begin{bmatrix} A^k & P_t(B) & 0 \\ 0 & A^tBD^t & P_t(B) \\ 0 & 0 & N^l \end{bmatrix} - r\begin{bmatrix} A^k & P_t(B) \\ 0 & N^l \end{bmatrix}.$$

Thus we have the desired results. □

Theorem 14.3. *Let M be given by* (14.1) *with* $\mathrm{Ind}(M) = 1$. *Then the ranks of G_1—G_4 in the group inverse of M in* (14.3) *can be expressed as*

$$r(G_1) = r\begin{bmatrix} V_2DW_2 & V_1 \\ W_1 & 0 \end{bmatrix} - r(M), \qquad r(G_2) = r\begin{bmatrix} V_1BW_2 & V_2 \\ W_1 & 0 \end{bmatrix} - r(M), \tag{14.13}$$

$$r(G_3) = r\begin{bmatrix} V_2CW_1 & V_1 \\ W_2 & 0 \end{bmatrix} - r(M), \qquad r(G_4) = r\begin{bmatrix} V_1AW_1 & V_2 \\ W_2 & 0 \end{bmatrix} - r(M), \tag{14.14}$$

where V_1, V_2, W_1 and W_2 are defined in (14.2).

Proof. Note that $M^{\#} = M(M^3)^{\dagger}M$ when $\mathrm{Ind}(M) = 1$. Thus G_1 in (14.13) can be written as $G_1 = W_1(M^3)^{\dagger}V_1$. In that case it follows by (1.7) that

$$\begin{aligned} r(G_1) &= r\begin{bmatrix} M^3 & V_1 \\ W_1 & 0 \end{bmatrix} - r(M^3) \\ &= r\begin{bmatrix} [\,0,\ V_2\,]M\begin{bmatrix} 0 \\ W_2 \end{bmatrix} & V_1 \\ W_1 & 0 \end{bmatrix} - r(M) = r\begin{bmatrix} V_2DW_2 & V_1 \\ W_1 & 0 \end{bmatrix} - r(M). \end{aligned}$$

In the same manner we can show the other three in (14.13) and (14.14). □

Corollary 14.4. *Let M be given by* (14.1) *with* $\mathrm{Ind}(M) = 1$.

(a) *If M satisfies the rank additivity condition*

$$r(M) = r(V_1) + r(V_2) = r(W_1) + r(W_2),$$

then the ranks of G_1—G_4 in the group inverse of M in (14.3) *can be expressed as*

$$\begin{aligned} r(G_1) &= r(V_1) + r(W_1) + r(V_2DW_2) - r(M), \\ r(G_2) &= r(V_2) + r(W_1) + r(V_1BW_2) - r(M), \\ r(G_3) &= r(V_1) + r(W_2) + r(V_2CW_1) - r(M), \\ r(G_4) &= r(V_2) + r(W_2) + r(V_1AW_1) - r(M). \end{aligned}$$

(b) *If M satisfies the rank additivity condition*

$$r(M) = r(A) + r(B) + r(C) + r(D),$$

then the ranks of G_1—G_4 in the group inverse of M in (14.3) *satisfy*

$$r(G_1) = r(A) - r(D) + r(V_2DW_2), \qquad r(G_2) = r(B) - r(C) + r(V_1BW_2),$$
$$r(G_3) = r(C) - r(B) + r(V_2CW_1), \qquad r(G_2) = r(D) - r(A) + r(V_1AW_1).$$

where V_1, V_2, W_1 and W_2 are defined in (14.2).

In addition, we have some inequalities on ranks of submatrices in the group inverse of a block matrix.

Corollary 14.5. *Let M be given by* (14.1) *with* $\mathrm{Ind}(M) = 1$. *Then the ranks of the matrices G_1—G_4 in* (14.3) *satisfy the following rank inequalities*

(a) $r(G_1) \geq r(V_1) + r(W_1) - r(M)$.
(b) $r(G_1) \leq r(V_1) + r(W_1) + r(D) - r(M)$.
(c) $r(G_2) \geq r(V_2) + r(W_1) - r(M)$.
(d) $r(G_2) \leq r(V_2) + r(W_1) + r(B) - r(M)$.
(e) $r(G_3) \geq r(V_1) + r(W_2) - r(M)$.
(f) $r(G_3) \leq r(V_1) + r(W_2) + r(C) - r(M)$.
(g) $r(G_4) \geq r(V_2) + r(W_2) - r(M)$.
(h) $r(G_4) \leq r(V_2) + r(W_2) + r(A) - r(M)$.

Proof. Follows from (14.13) and (14.14). □.

Chapter 15

Reverse order laws for Drazin inverses

In this chapter we consider reverse order laws for Drazin inverses of products of matrices. We will give necessary and sufficient conditions for $(ABC)^D = C^D B^D A^D$ to hold and then present some of its consequences.

Lemma 15.1. *Let A, $X \in \mathcal{C}^{m\times m}$ with $\operatorname{Ind}(A) = k$. Then $X = A^D$ if and only if*

$$A^{k+1}X = A^k, \qquad XA^{k+1} = A^k, \qquad \text{and} \qquad r(X) = r(A^k). \tag{15.1}$$

Proof. Follows from the definition of the Drazin inverse of a matrix. $\square$

Lemma 15.2. *Let A, B, $C \in \mathcal{C}^{m\times m}$ with $\operatorname{Ind}(A) = k_1$, $\operatorname{Ind}(B) = k_2$ and $\operatorname{Ind}(C) = k_3$. Then the product $C^D B^D A^D$ of the Drazin inverses of A, B, and C can be expressed in the form*

$$C^D B^D A^D = [\, C^{k_3},\ 0,\ 0\,] \begin{bmatrix} 0 & 0 & A^{2k_1+1} \\ 0 & B^{2k_2+1} & B^{k_2}A^{k_1} \\ C^{2k_3+1} & C^{k_3}B^{k_2} & 0 \end{bmatrix}^{\dagger} \begin{bmatrix} A^{k_1} \\ 0 \\ 0 \end{bmatrix} := PN^{\dagger}Q, \tag{15.2}$$

where P, N and Q satisfy the three properties

$$R(Q) \subseteq R(N), \qquad R(P^*) \subseteq R(N^*), \qquad r(N) = r(A^{k_1}) + r(B^{k_2}) + r(C^{k_3}). \tag{15.3}$$

Proof. It is easy to verify that the 3×3 block matrix N in (15.2) satisfies the conditions in Lemma 8.8. Hence it follows by (8.8) that

$$N^{\dagger} = \begin{bmatrix} (C^{2k_3+1})^{\dagger}C^{k_3}B^{k_2}(B^{2k_2+1})^{\dagger}B^{k_2}A^{k_1}(A^{2k_1+1})^{\dagger}A^{k_1} & * & * \\ * & * & 0 \\ * & 0 & 0 \end{bmatrix}. \tag{15.4}$$

Thus we have (15.2). The three properties in (15.3) follows from the structure of N. $\square$

The main results of the chapter are the following two.

Theorem 15.3. *Let A, B, $C \in \mathcal{C}^{m\times m}$ with $\operatorname{Ind}(A) = k_1$, $\operatorname{Ind}(B) = k_2$ and $\operatorname{Ind}(C) = k_3$, and denote $M = ABC$ with $\operatorname{Ind}(M) = t$. Then the reverse order law $(ABC)^D = C^D B^D A^D$ holds if and only if A, B and C satisfy the three rank equalities*

$$r \begin{bmatrix} 0 & 0 & A^{2k_1+1} & A^{k_1} \\ 0 & B^{2k_2+1} & B^{k_2}A^{k_1} & 0 \\ C^{2k_3+1} & C^{k_3}B^{k_2} & 0 & 0 \\ M^{t+1}C^{k_3} & & 0 & M^t \end{bmatrix} = r(A^{k_1}) + r(B^{k_2}) + r(C^{k_3}), \tag{15.5}$$

$$r\begin{bmatrix} 0 & 0 & A^{2k_1+1} & A^{k_1}M^{t+1} \\ 0 & B^{2k_2+1} & B^{k_2}A^{k_1} & 0 \\ C^{2k_3+1} & C^{k_3}B^{k_2} & 0 & 0 \\ C^{k_3} & & 0 & M^t \end{bmatrix} = r(A^{k_1}) + r(B^{k_2}) + r(C^{k_3}), \tag{15.6}$$

$$r\begin{bmatrix} B^{2k_2+1} & B^{k_2}A^{k_1} \\ C^{k_3}B^{k_2} & 0 \end{bmatrix} = r(B^{k_2}) + r(M^t). \tag{15.7}$$

Proof. Let $X = C^D B^D A^D$. Then by definition of the Drazin inverse, $X = M^D$ if and only if $M^{t+1}X = M^t$, $XM^{t+1} = M^t$ and $r(X) = r(M^t)$, which are equivalent to

$$r(\, M^k - M^{k+1}X\,) = 0, \qquad r(\, M^k - XM^{k+1}\,) = 0 \;\text{ and }\; r(X) = r(M^t). \tag{15.8}$$

Replacing X in (15.8) by $X = PN^{\dagger}Q$ in (15.2) and applying (1.7) them, we find that

$$r(\, M^t - M^{t+1}X\,) = r(\, M^t - M^{t+1}PN^{\dagger}Q\,) = r\begin{bmatrix} N & Q \\ M^{t+1}P & M^t \end{bmatrix} - r(N),$$

$$r(\, M^t - XM^{t+1}\,) = r(\, M^t - PN^{\dagger}QM^{t+1}\,) = r\begin{bmatrix} N & QM^{t+1} \\ P & M^t \end{bmatrix} - r(N),$$

$$r(X) = r(PN^{\dagger}Q) = r\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix} - r(N).$$

Putting them in (15.8), we obtain (15.5)—(15.7). □

Theorem 15.4. *Let $A, B, C \in \mathcal{C}^{m\times m}$ with* $\mathrm{Ind}(A) = k_1$, $\mathrm{Ind}(B) = k_2$ *and* $\mathrm{Ind}(C) = k_3$, *and let $M = ABC$ with* $\mathrm{Ind}(M) = t$. *Then the reverse order law $(ABC)^D = C^D B^D A^D$ holds if and only if A, B and C satisfy the following rank equality*

$$r\begin{bmatrix} 0 & 0 & A^{2k_1+1} & A^{k_1} & 0 \\ 0 & B^{2k_2+1} & B^{k_2}A^{k_1} & 0 & 0 \\ C^{2k_3+1} & C^{k_3}B^{k_2} & 0 & 0 & 0 \\ C^{k_3} & 0 & 0 & 0 & M^t \\ 0 & 0 & 0 & M^t & M^{2t+1} \end{bmatrix} = r(A^{k_1}) + r(B^{k_2}) + r(C^{k_3}) - r(M^t). \tag{15.9}$$

Proof. Applying (2.3) to $(ABC)^D - C^D B^D A^D = M^t(M^{2t+1})^{\dagger}M^t - PN^{\dagger}Q$, we find that

$$\begin{aligned} r[\,(ABC)^D - C^D B^D A^D\,] &= r[\,PN^{\dagger}Q - M^t(M^{2t+1})^{\dagger}M^t\,] \\ &= r\begin{bmatrix} N & 0 & Q \\ 0 & -M^{2t+1} & M^t \\ P & M^t & 0 \end{bmatrix} - r(N) - r(M^t) \\ &= r\begin{bmatrix} N & 0 & Q \\ 0 & 0 & M^t \\ P & M^t & -M^{2t+1} \end{bmatrix} - r(N) - r(M^t). \end{aligned}$$

Thus (15.9) follows by putting P, N and Q in it. □

We next give some particular cases of the above two theorems.

Corollary 15.5. *Let $A, B, C \in \mathcal{C}^{m\times m}$ with* $\mathrm{Ind}(B) = k$ *and* $\mathrm{Ind}(ABC) = t$, *where A and C are nonsingular. Then*

(a) $r[\,(ABC)^D - C^{-1}B^DA^{-1}\,] = r\begin{bmatrix} B^k \\ (ABC)^tA \end{bmatrix} + r[\,B^k,\; C(ABC)^t\,] - r(B^k) - r[(ABC)^t]$.

(b) $(ABC)^D = C^{-1}B^DA^{-1} \Leftrightarrow R[C(ABC)^t] = R(B^k)$ *and* $R\{[(ABC)^tA]^*\} = R[(B^k)^*]$.

Proof. It is easy to verify that both $(ABC)^D$ and $C^{-1}B^DA^{-1}$ are outer inverses of ABC. Thus it follows from (5.1) that

$$\begin{aligned} &r[\,(ABC)^D - C^{-1}B^DA^{-1}\,] \\ &= r\begin{bmatrix} (ABC)^D \\ C^{-1}B^DA^{-1} \end{bmatrix} + r[\,(ABC)^D,\; C^{-1}B^DA^{-1}\,] - r[(ABC)^D] - r(B^D) \\ &= r\begin{bmatrix} (ABC)^tA \\ B^k \end{bmatrix} + r[\,C(ABC)^t,\; B^k\,] - r[(ABC)^t] - r(B^k), \end{aligned}$$

as required for Part (a). Notice that

$$r\begin{bmatrix} B^k \\ (ABC)^tA \end{bmatrix} \geq r(B^k), \quad r\begin{bmatrix} B^k \\ (ABC)^tA \end{bmatrix} \geq r[(ABC)^t],$$

and

$$r[\, B^l, \ C(ABC)^t\,] \geq r(B^k), \quad r[\, B^l, \ C(ABC)^t\,] \geq r[(ABC)^t].$$

Then Part (b) follows from Part (a). $\square$

Corollary 15.6. *Let* A, B, $C \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k_1$, $\mathrm{Ind}(B) = k_2$ *and* $\mathrm{Ind}(C) = k_3$, *and let* $M = ABC$ *with* $\mathrm{Ind}(M) = t$. *Moreover suppose that*

$$AB = BA, \qquad AC = CA, \qquad BC = CB. \tag{15.10}$$

Then the reverse order law $(ABC)^D = C^DB^DA^D$ *holds if and only if* A, B *and* C *satisfy* (15.7).

Proof. It is not difficult to verify that under (15.10), the two rank equalities in (15.5) and (15.6) become two identities. Thus, (15.7) becomes a necessary and sufficient condition for $(ABC)^D = C^DB^DA^D$ to hold. $\square$

Corollary 15.7. *Let* A, $B \in \mathcal{C}^{m\times m}$ *with* $\mathrm{Ind}(A) = k$, $\mathrm{Ind}(B) = l$ *and* $\mathrm{Ind}(AB) = t$. *Then the following three are equivalent*:

(a) $(AB)^D = B^DA^D$.

(b) $r\begin{bmatrix} 0 & A^{2k+1} & A^k & 0 \\ B^{2l+1} & B^lA^k & 0 & 0 \\ B^l & 0 & 0 & (AB)^t \\ 0 & 0 & (AB)^t & (AB)^{2t+1} \end{bmatrix} = r(A^k) + r(B^l) - r[(AB)^t].$

(c) *The following three rank equalities are all satisfied*

$$r[(AB)^t] = r(B^lA^k),$$

$$r\begin{bmatrix} 0 & A^{2k+1} & A^k \\ B^{2l+1} & B^lA^k & 0 \\ (AB)^{t+1}B^l & 0 & -(AB)^t \end{bmatrix} = r(A^k) + r(B^l),$$

$$r\begin{bmatrix} 0 & A^{2k+1} & A^k(AB)^{t+1} \\ B^{2l+1} & B^lA^k & 0 \\ B^l & 0 & -(AB)^t \end{bmatrix} = r(A^k) + r(B^l).$$

Proof. Letting $C = I_m$ in (15.9) results in Part (b), and letting $B = I_m$ and replacing C by B in Theorem 15.4 result in Part (c). $\square$

Chapter 16

Ranks equalities for weighted Moore-Penrose inverses

The weighted Moore-Penrose inverse of a matrix $A \in \mathcal{C}^{m\times n}$ with respect to two positive definite matrices $M \in \mathcal{C}^{m\times m}$ and $N \in \mathcal{C}^{n\times n}$ is defined to be the unique solution of the following four matrix equations

$$AXA = A, \qquad XAX = X, \qquad (MAX)^* = MAX, \qquad (NXA)^* = NXA, \tag{16.1}$$

and this X is often denoted by $X = A^{\dagger}_{M,N}$. In particular, when $M = I_m$ and $N = I_n$, $A^{\dagger}_{M,N}$ is the standard Moore-Penrose inverse $A^{\dagger}$ of A. As is well known (see, e.g., Rao and Mitra [118]), the weighted Moore-Penrose inverse $A^{\dagger}_{M,N}$ of A can be written as a matrix expressions involving a standard Moore-Penrose inverse as follows

$$A^{\dagger}_{M,N} = N^{-\frac{1}{2}}(M^{\frac{1}{2}}AN^{-\frac{1}{2}})^{\dagger}M^{\frac{1}{2}}, \tag{16.2}$$

where $M^{\frac{1}{2}}$ and $N^{\frac{1}{2}}$ are the positive definite square roots of M and N, respectively. According to (16.2), it is easy to verify that

$$R(A^{\dagger}_{M,N}) = R(N^{-1}A^*), \quad \text{and} \quad R[(A^{\dagger}_{M,N})^*] = R(MA). \tag{16.2}$$

Based on these basic facts and the rank formulas in Chapters 2—5, we now can establish various rank equalities related to weighted Moore-Penrose inverses of matrices, and the consider their various consequences.

Theorem 16.1. *Let $A \in \mathcal{C}^{m\times n}$ be given, $M \in \mathcal{C}^{m\times m}$ and $N \in \mathcal{C}^{n\times n}$ be two positive definite matrices. Then*

(a) $r(\, A^{\dagger} - A^{\dagger}_{M,N}\,) = r\begin{bmatrix} A \\ AN \end{bmatrix} + r[\, A, \ MA\,] - 2r(A).$

(b) $r(\, A^{\dagger} - A^{\dagger}_{M,I}\,) = r[\, A, \ MA\,] - r(A).$

(c) $r(\, A^{\dagger} - A^{\dagger}_{I,N}\,) = r\begin{bmatrix} A \\ AN \end{bmatrix} - r(A).$

(d) $A^{\dagger}_{M,N} = A^{\dagger} \Leftrightarrow R(MA) = R(A)$ *and* $R[(AN)^*] = R(A^*).$

Proof. Note that $A^{\dagger}$ and $A^{\dagger}_{M,N}$ are outer inverses of A. Thus it follows from (5.1) that

$$\begin{aligned} r(\, A^{\dagger} - A^{\dagger}_{M,N}\,) &= r\begin{bmatrix} A^{\dagger} \\ A^{\dagger}_{M,N} \end{bmatrix} + r[\, A^{\dagger}, \ A^{\dagger}_{M,N}\,] - r(A^{\dagger}) - r(A^{\dagger}_{M,N}) \\ &= r\begin{bmatrix} A^* \\ (MA)^* \end{bmatrix} + r[\, A^*, \ N^{-1}A^*\,] - 2r(A) \\ &= r\begin{bmatrix} A \\ AN \end{bmatrix} + r[\, A, \ MA\,] - 2r(A). \end{aligned}$$

Parts (a)—(c) follow immediately from it. $\square$

Theorem 16.2. *Let $A \in \mathcal{C}^{m\times n}$ be given, $M \in \mathcal{C}^{m\times m}$ and $N \in \mathcal{C}^{n\times n}$ be two positive definite matrices. Then*

(a) $r(\, AA^{\dagger}_{M,N} - AA^{\dagger}\,) = r[\, A,\ MA\,] - r(A)$.

(b) $r(\, A^{\dagger}_{M,N}A - A^{\dagger}A\,) = r\begin{bmatrix} A \\ AN \end{bmatrix} - r(A)$.

(c) $AA^{\dagger}_{M,N} = AA^{\dagger} \Leftrightarrow R(MA) = R(A)$.

(d) $A^{\dagger}_{M,N}A = A^{\dagger}A \Leftrightarrow R[(AN)^*] = R(A^*)$.

Proof. Note that both $AA^{\dagger}$ and $AA^{\dagger}_{M,N}$ are idempotent. It follows from (3.1) that

$$\begin{aligned} r(\, AA^{\dagger} - AA^{\dagger}_{M,N}\,) &= r\begin{bmatrix} AA^{\dagger} \\ AA^{\dagger}_{M,N} \end{bmatrix} + r[\, AA^{\dagger},\ AA^{\dagger}_{M,N}\,] - r(AA^{\dagger}) - r(AA^{\dagger}_{M,N}) \\ &= r\begin{bmatrix} A^* \\ (MA)^* \end{bmatrix} + r[\, A,\ A\,] - 2r(A) \\ &= r[\, A,\ MA\,] - r(A), \end{aligned}$$

as required for Part (a). Similarly we can show Part (b). □

Theorem 16.3. *Let $A \in \mathcal{C}^{m\times m}$ be given, and $M,\, N \in \mathcal{C}^{m\times m}$ be two positive definite matrices. Then*

(a) $r(\, AA^{\dagger}_{M,N} - A^{\dagger}_{M,N}A\,) = r[\, A^*,\ MA\,] + r[\, A^*,\ NA\,] - 2r(A)$.

(b) $AA^{\dagger}_{M,N} = A^{\dagger}_{M,N}A \Leftrightarrow R(MA) = R(NA) = R(A^*) \Leftrightarrow$ *both MA and NA are EP.*

Proof. Note that both $AA^{\dagger}$ and $AA^{\dagger}_{M,N}$ are idempotent. It follows by (3.1) that

$$\begin{aligned} r(\, AA^{\dagger}_{M,N} - A^{\dagger}_{M,N}A\,) &= r\begin{bmatrix} AA^{\dagger}_{M,N} \\ A^{\dagger}_{M,N}A \end{bmatrix} + r[\, AA^{\dagger}_{M,N},\ A^{\dagger}_{M,N}A\,] - r(AA^{\dagger}_{M,N}) - r(A^{\dagger}_{M,N}A) \\ &= r\begin{bmatrix} A^{\dagger}_{M,N} \\ A \end{bmatrix} + r[\, A,\ A^{\dagger}_{M,N}\,] - 2r(A) \\ &= r\begin{bmatrix} (MA)^* \\ A \end{bmatrix} + r[\, A,\ N^{-1}A^*\,] - 2r(A), \end{aligned}$$

as required for Part (a). Part(b) follows immediately from Part (a). □

Based on the result in Theorem 16.3(b), we can extend the concept of EP matrix to weighted case: A square matrix A is said to be *weighted EP* if both MA and NA are EP, where both M and N are two positive definite matrices. It is expected that weighted EP matrix would have some nice properties. But we do not intend to go further along this direction in the thesis.

Theorem 16.4. *Let $A \in \mathcal{C}^{m\times m}$ be given, and $M,\, N \in \mathcal{C}^{m\times m}$ be two positive definite matrices. Then*

(a) $r(\, AA^{\dagger}_{M,N} - \overline{A^{\dagger}_{M,N}A}\,) = r[\, A^T,\ MA\,] + r[\, A^T,\ N^TA\,] - 2r(A)$.

(b) $AA^{\dagger}_{M,N} = \overline{A^{\dagger}_{M,N}A} \Leftrightarrow\ R(MA) = R(N^TA) = R(A^T) \Leftrightarrow$ *both MA and N^TA are EP.*

Proof. Follows from (3.1) by noting that both $AA^{\dagger}_{M,N}$ and $\overline{AA^{\dagger}_{M,N}}$ are idempotent. □

Theorem 16.5. *Let $A \in \mathcal{C}^{m\times m}$ be given with* $\mathrm{Ind}(A) = 1$, *and $M,\, N \in \mathcal{C}^{m\times m}$ be two positive definite matrices. Then*

(a) $r(\, A^{\dagger}_{M,N} - A^{\#}\,) = r[\, A^*,\ MA\,] + r[\, A^*,\ NA\,] - 2r(A)$.

(b) $A^{\dagger}_{M,N} = A^{\#} \Leftrightarrow R(MA) = R(NA) = R(A^*)$, *i.e., A is weighted EP.*

Proof. Note that both $A^{\dagger}$ and $A^{\#}$ are outer inverses of A. It follows by (5.1) that

$$\begin{aligned} r(\, A^{\dagger}_{M,N} - A^{\#}\,) &= r\begin{bmatrix} A^{\dagger}_{M,N} \\ A^{\#} \end{bmatrix} + r[\, A^{\dagger}_{M,N},\ A^{\#}\,] - r(A^{\dagger}_{M,N}) - r(A^{\#}) \\ &= r\begin{bmatrix} (MA)^* \\ A \end{bmatrix} + r[\, N^{-1}A^*,\ A\,] - 2r(A) \\ &= r[\, A^*,\ MA\,] + r[\, A^*,\ NA\,] - 2r(A), \end{aligned}$$

as required for Part (a). □

Theorem 16.6. *Let $A \in \mathcal{C}^{m\times m}$ be given with* $\mathrm{Ind}(A) = 1$, *and* $M, N \in \mathcal{C}^{m\times m}$ *be two positive definite matrices. Then*

(a) $r(AA^{\dagger}_{M,N} - AA^{\#}) = r[A^*, \ MA] - r(A)$.

(b) $r(A^{\dagger}_{M,N}A - A^{\#}A) = r[A^*, \ NA] - r(A)$.

(c) $r(A^{\dagger}_{M,N} - A^{\#}) = r(AA^{\dagger}_{M,N} - AA^{\#}) + r(A^{\dagger}_{M,N}A - A^{\#}A)$.

In particular,

(d) $AA^{\dagger}_{M,N} = AA^{\#} \Leftrightarrow R(MA) = R(A^*)$, *i.e.,* MA *is EP.*

(e) $A^{\dagger}_{M,N}A = A^{\#}A \Leftrightarrow R(NA) = R(A^*)$, *i.e.,* NA *is EP.*

(f) $A^{\dagger}_{M,N} = A^{\#} \Leftrightarrow AA^{\dagger}_{M,N} = AA^{\#}$ *and* $A^{\dagger}_{M,N}A = A^{\#}A$.

Proof. Note that both $AA^{\dagger}$ and $AA^{\#}$ are idempotent. It follows from (5.1) that

$$\begin{aligned} r(AA^{\dagger}_{M,N} - AA^{\#}) &= r\begin{bmatrix} AA^{\dagger}_{M,N} \\ AA^{\#} \end{bmatrix} + r[AA^{\dagger}_{M,N}, \ AA^{\#}] - r(AA^{\dagger}_{M,N}) - r(AA^{\#}) \\ &= r\begin{bmatrix} A^{\dagger}_{M,N} \\ A^{\#} \end{bmatrix} + r[A, \ A] - 2r(A) \\ &= r\begin{bmatrix} (MA)^* \\ A^* \end{bmatrix} = r[A^*, \ MA] - r(A), \end{aligned}$$

as required for Part (a). □

Theorem 16.7. *Let $A \in \mathcal{C}^{m\times m}$ be given with* $\mathrm{Ind}(A) = k$, *and* $M, N \in \mathcal{C}^{m\times m}$ *be two positive definite matrices. Then*

(a) $r(A^{\dagger}_{M,N} - A^{D}) = r\begin{bmatrix} A^kM^{-1} \\ A^* \end{bmatrix} + r[NA^k, \ A^*] - r(A) - r(A^k)$.

(b) $r(A^{\dagger}_{M,N} - A^{D}) = r(A^{\dagger}_{M,N}) - r(A^{D}) \Leftrightarrow R(NA^k) \subseteq r(A^*)$ *and* $R[(A^kM^{-1})^*] \subseteq r(A)$.

Proof. Note that both $A^{\dagger}_{M,N}$ and A^{D} are outer inverses of A. It follows by (5.1) that

$$\begin{aligned} r(A^{\dagger}_{M,N} - A^{D}) &= r\begin{bmatrix} A^{\dagger}_{M,N} \\ A^{D} \end{bmatrix} + r[A^{\dagger}_{M,N}, \ A^{D}] - r(A^{\dagger}_{M,N}) - r(A^{D}) \\ &= r\begin{bmatrix} (MA)^* \\ A^k \end{bmatrix} + r[N^{-1}A^*, \ A^k] - 2r(A) \\ &= r\begin{bmatrix} A^* \\ A^kM^{-1} \end{bmatrix} + r[A^*, \ NA^k] - 2r(A), \end{aligned}$$

as required for Part (a). □

Theorem 16.8. *Let $A \in \mathcal{C}^{m\times m}$ be given with* $\mathrm{Ind}(A) = k$, *and* $M, N \in \mathcal{C}^{m\times m}$ *be two positive definite matrices. Then*

(a) $r(AA^{\dagger}_{M,N} - AA^{D}) = r\begin{bmatrix} A^kM^{-1} \\ A^* \end{bmatrix} - r(A^k)$.

(b) $r(A^{\dagger}_{M,N}A - A^{D}A) = r[NA^k, \ A^*] - r(A^k)$.

(c) $r(A^{\dagger}_{M,N} - A^{D}) = r(AA^{\dagger}_{M,N} - AA^{D}) + (A^{\dagger}_{M,N}A - A^{D}A) + r(A^k) - r(A)$.

Proof. Note that both $AA^{\dagger}_{M,N}$ and AA^{D} are idempotent. It follows from (5.1) that

$$\begin{aligned} r(AA^{\dagger}_{M,N} - AA^{D}) &= r\begin{bmatrix} AA^{\dagger}_{M,N} \\ AA^{D} \end{bmatrix} + r[AA^{\dagger}_{M,N}, \ AA^{D}] - r(AA^{\dagger}_{M,N}) - r(AA^{D}) \\ &= r\begin{bmatrix} A^{\dagger}_{M,N} \\ A^{D} \end{bmatrix} + r[A, \ A^{D}] - r(A) - r(A^k) \\ &= r\begin{bmatrix} (MA)^* \\ A^k \end{bmatrix} - r(A) = r\begin{bmatrix} A^* \\ A^kM^{-1} \end{bmatrix} - r(A), \end{aligned}$$

as required for Part (a). Similarly we can show Part (b). Combining Theorem 16.6(a) and Theorem 16.7(a) yields Part (c). $\square$

Theorem 16.9. *Let $A \in \mathcal{C}^{m\times n}$ be given, $M,\ N \in \mathcal{C}^{m\times m}$ be two positive definite matrices. Then*

(a) $r(\, A^{\dagger}_{M,N}A^k - A^kA^{\dagger}_{M,N}\,) = r\begin{bmatrix} A^k \\ A^*M \end{bmatrix} + r[\, A^k,\ N^{-1}A^*\,] - 2r(A).$

(b) $A^{\dagger}_{M,N}A^k = A^kA^{\dagger}_{M,N} \Leftrightarrow R(A^k) \subseteq R(N^{-1}A^*)$ *and* $R[(A^k)^*] \subseteq R(MA).$

Proof. Follows from (4.1). $\square$

Based on the result in Theorem 16.9(b), we can extend the concept of power-EP matrix to weighted case: A square matrix A is said to be *weighted power-EP* if both $R(A^k) \subseteq R(N^{-1}A^*)$ and $R[(A^k)^*] \subseteq R(MA)$ hold, where both M and N are positive definite matrices.

Theorem 16.10. *Let $A \in \mathcal{C}^{m\times m}$ be given with* $\mathrm{Ind}(A) = k$, *and $M,\ N \in \mathcal{C}^{m\times m}$ be two positive definite matrices. Then*

(a) $r(\, A^{\dagger}_{M,N}A^D - A^DA^{\dagger}_{M,N}\,) = r\begin{bmatrix} A^k \\ A^*M \end{bmatrix} + r[\, A^k,\ N^{-1}A^*\,] - 2r(A) = r(\, A^{\dagger}_{M,N}A^k - A^kA^{\dagger}_{M,N}\,).$

(b) $A^{\dagger}_{M,N}A^D = A^DA^{\dagger}_{M,N} \Leftrightarrow R(A^k) \subseteq R(N^{-1}A^*)$ *and* $R[(A^k)^*] \subseteq R(MA)$, *i.e., A is weighted power-EP.*

Proof. Follows from (4.1). $\square$

Theorem 16.11. *Let $A \in \mathcal{C}^{m\times n}$ be given, $M,\ S \in \mathcal{C}^{m\times m}$ and $N,\, T \in \mathcal{C}^{n\times n}$ be four positive definite matrices. Then*

(a) $r(\, A^{\dagger}_{M,N} - A^{\dagger}_{S,T}\,) = r\begin{bmatrix} AN^{-1} \\ AT^{-1} \end{bmatrix} + r[\, MA,\ SA\,] - 2r(A).$

(b) $A^{\dagger}_{M,N} = A^{\dagger}_{S,T} \Leftrightarrow R(MA) = R(SA)$ *and* $R[(AN^{-1})^*] = R[(AT^{-1})^*].$

Proof. Note that both $A^{\dagger}_{M,N}$ and $A^{\dagger}_{P,Q}$ are outer inverses of A. Thus it follows by Eq.(5.1) that

$$\begin{aligned} r(\, A^{\dagger}_{M,N} - A^{\dagger}_{S,T}\,) &= r\begin{bmatrix} A^{\dagger}_{M,N} \\ A^{\dagger}_{S,T} \end{bmatrix} + r[\, A^{\dagger}_{M,N},\ A^{\dagger}_{S,T}\,] - r(A^{\dagger}_{M,N}) - r(A^{\dagger}_{S,T}) \\ &= r\begin{bmatrix} (MA)^* \\ (SA)^* \end{bmatrix} + r[\, N^{-1}A^*,\ T^{-1}A^*\,] - 2r(A) \\ &= r\begin{bmatrix} AN^{-1} \\ AT^{-1} \end{bmatrix} + r[\, MA,\ SA\,] - 2r(A), \end{aligned}$$

establishing Part (a). $\square$

Theorem 16.12. *Let $A \in \mathcal{C}^{m\times n}$ be given, $M,\ S \in \mathcal{C}^{m\times m}$, $N,\ T \in \mathcal{C}^{n\times n}$ be four positive definite matrices. Then*

(a) $r(\, AA^{\dagger}_{M,N} - AA^{\dagger}_{S,T}\,) = r[\, MA,\ SA\,] - 2r(A).$

(b) $r(\, A^{\dagger}_{M,N}A - A^{\dagger}_{S,T}A\,) = r\begin{bmatrix} AN^{-1} \\ AT^{-1} \end{bmatrix} - r(A).$

(c) $r(\, A^{\dagger}_{M,N} - A^{\dagger}_{S,T}\,) = r(\, AA^{\dagger}_{M,N} - AA^{\dagger}_{S,T}\,) + r(\, A^{\dagger}_{M,N}A - A^{\dagger}_{S,T}A\,).$

Proof. Follows from (3.1) by noticing that $AA^{\dagger}_{M,N}$, $A^{\dagger}_{M,N}A$, $AA^{\dagger}_{S,T}$ and $A^{\dagger}_{S,T}A$ are idempotent matrices. $\square$

Theorem 16.13. *Let $A \in \mathcal{C}^{m\times m}$ be an idempotent or tripotent matrix, and $M,\ N \in \mathcal{C}^{m\times m}$ be two positive definite matrices. Then*

(a) $r(\, A - A^{\dagger}_{M,N}\,) = r[\, A^*,\ MA\,] + r[\, A^*,\ NA\,] - 2r(A).$

(b) $A = A^{\dagger}_{M,N} \Leftrightarrow R(MA) = R(NA) = R(A^*)$, *i.e., A is weighted EP.*

Proof. Note that $A,\ A^{\dagger}_{M,N} \in A\{2\}$ when A is idempotent or tripotent. It follows by (5.1) that

$$r(\, A - A^{\dagger}_{M,N}\,) \ = \ r\begin{bmatrix} A \\ A^{\dagger}_{M,N} \end{bmatrix} + r[\, A,\ A^{\dagger}_{M,N}\,] - r(A) - r(A^{\dagger}_{M,N})$$

$$= r\begin{bmatrix} A \\ (MA)^* \end{bmatrix} + r[\,A,\ N^{-1}A^*\,] - 2r(A)$$
$$= r[\,A^*,\ MA\,] + r[\,A^*,\ NA\,] - r(A) - r(A^k),$$

as required for Part (a). □

Theorem 16.14. *Let $A,\, B \in \mathcal{C}^{m\times m}$ be given, $M,\, N \in \mathcal{C}^{m\times m}$ be two positive definite matrices. Then*

(a) $r(\, AA^{\dagger}_{M,N}B - BA^{\dagger}_{M,N}A\,) = r\begin{bmatrix} A \\ A^*MB \end{bmatrix} + r[\,A,\ BN^{-1}A^*\,] - 2r(A)$.

(b) $r(\, A^{\dagger}_{M,N}AB - BAA^{\dagger}_{M,N}\,) = r\begin{bmatrix} AB \\ A^*M \end{bmatrix} + r[\,BA,\ N^{-1}A^*\,] - 2r(A)$.

(c) $AA^{\dagger}_{M,N}B = BA^{\dagger}_{M,N}A \Leftrightarrow R(BN^{-1}A^*) \subseteq R(A)$ *and* $R(B^*MA) \subseteq R(A^*)$.

(d) $A^{\dagger}_{M,N}AB = BAA^{\dagger}_{M,N} \Leftrightarrow R(BA) \subseteq R(N^{-1}A^*)$ *and* $R[(AB)^*] \subseteq R(MA)$.

Proof. Parts (a) and (b) Follow from (4.1) by noticing that both $AA^{\dagger}_{M,N}$ and $A^{\dagger}_{M,N}A$ are idempotent. □

Theorem 16.15. *Let $A \in \mathcal{C}^{m\times m}$ be given with* $\mathrm{Ind}(A) = 1$, *and $P,\, Q \in \mathcal{C}^{m\times m}$ be two nonsingular matrices. Then*

(a) $r[\,(PAQ)^{\dagger} - Q^{-1}A^{\#}P^{-1}\,] = r\begin{bmatrix} A \\ A^*P^*P \end{bmatrix} + r[\,A,\ QQ^*A^*\,] - 2r(A)$.

(b) $(PAQ)^{\dagger} = Q^{-1}A^{\#}P^{-1} \Leftrightarrow R(QQ^*A^*) = R(A)$ *and* $R(P^*PA) = R(A^*)$.

Proof. It is easy to verify that both $(PAQ)^{\dagger}$ and $Q^{-1}A^{\#}P^{-1}$ are outer inverses of PAQ. Thus it follows by (5.1) that

$$\begin{aligned} & r[\,(PAQ)^{\dagger} - Q^{-1}A^{\#}P^{-1}\,] \\ &= r\begin{bmatrix} (PAQ)^{\dagger} \\ Q^{-1}A^{\#}P^{-1} \end{bmatrix} + r[\,(PAQ)^{\dagger},\ Q^{-1}A^{\#}P^{-1}\,] - r[(PAQ)^{\dagger}] - r[Q^{-1}A^{\#}P^{-1}] \\ &= r\begin{bmatrix} (PAQ)^*P \\ A^{\#} \end{bmatrix} + r[\,Q(PAQ)^*,\ A^{\#}\,] - 2r(A) \\ &= r\begin{bmatrix} A^*P^*P \\ A \end{bmatrix} + r[\,QQ^*A^*,\ A\,] - 2r(A), \end{aligned}$$

establishing Part (a) and then Part (a). □

Theorem 16.16. *Let $A \in \mathcal{C}^{m\times m}$ be given, $M,\ N \in \mathcal{C}^{m\times m}$ be two positive definite matrices, and $P,\, Q \in \mathcal{C}^{m\times m}$ be two nonsingular matrices. Then*

(a) $r[\,(PAQ)^{\dagger} - Q^{-1}A^{\dagger}_{M,N}P^{-1}\,] = r\begin{bmatrix} A \\ AQQ^*N \end{bmatrix} + r[\,A,\ M^{-1}P^*PA^*\,] - 2r(A)$.

(b) $(PAQ)^{\dagger} = Q^{-1}A^{\dagger}_{M,N}P^{-1} \Leftrightarrow R(M^{-1}P^*PA^*) = R(A)$ *and* $R(NQQ^*A^*) = R(A^*)$.

Proof. It is easy to verify that both $(PAQ)^{\dagger}$ and $Q^{-1}A^{\dagger}_{M,N}P^{-1}$ are outer inverses of PAQ. Thus it follows by (5.1) that

$$\begin{aligned} & r[\,(PAQ)^{\dagger} - Q^{-1}A^{\dagger}_{M,N}P^{-1}\,] \\ &= r\begin{bmatrix} (PAQ)^{\dagger} \\ Q^{-1}A^{\dagger}_{M,N}P^{-1} \end{bmatrix} + r[\,(PAQ)^{\dagger},\ Q^{-1}A^{\dagger}_{M,N}P^{-1}\,] - r[(PAQ)^{\dagger}] - r[Q^{-1}A^{\dagger}_{M,N}P^{-1}] \\ &= r\begin{bmatrix} (PAQ)^*P \\ A^{\dagger}_{M,N} \end{bmatrix} + r[\,Q(PAQ)^*,\ A^{\dagger}_{M,N}\,] - 2r(A) \\ &= r\begin{bmatrix} A^*P^*P \\ (MA)^* \end{bmatrix} + r[\,QQ^*A^*,\ N^{-1}A^*\,] - 2r(A) \\ &= r\begin{bmatrix} A \\ AQQ^*N \end{bmatrix} + r[\,A,\ M^{-1}P^*PA\,] - 2r(A), \end{aligned}$$

establishing Part (a). □

Chapter 17

Reverse order laws for weighted Moore-Penrose inverses

Just as for Moore-Penrose inverses and Drazin inverses of products of matrices, we can also consider reverse order laws for weighted Moore-Penrose inverses of products of matrices. Noticing the basic fact in (16.2), we can easily extend the results in Chapter 8 to weighted Moore-Penrose inverses of products of matrices.

Theorem 17.1. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times k}$, and $C \in \mathcal{C}^{k\times l}$ be given and let $J = ABC$. Let $M \in \mathcal{C}^{m\times m}$, $N \in \mathcal{C}^{l\times l}$, $P \in \mathcal{C}^{n\times n}$, and $Q \in \mathcal{C}^{k\times k}$ be four positive definite matrices. Then the following three statements are equivalent*:

(a) $(ABC)^{\dagger}_{M,N} = C^{\dagger}_{Q,N}B^{\dagger}_{P,Q}A^{\dagger}_{M,P}$.

(b) $(M^{\frac{1}{2}}ABCN^{-\frac{1}{2}})^{\dagger} = (Q^{\frac{1}{2}}CN^{-\frac{1}{2}})^{\dagger}(P^{\frac{1}{2}}BQ^{-\frac{1}{2}})^{\dagger}(M^{\frac{1}{2}}AP^{-\frac{1}{2}})^{\dagger}$.

(c) $r\begin{bmatrix} BQ^{-1}B^*PB & 0 & BC \\ 0 & -JN^{-1}J^*MJ & JN^{-1}C^*QC \\ AB & AP^{-1}A^*MJ & 0 \end{bmatrix} = r(B) + r(J)$.

Proof. The equivalence of Part (a) and Part (b) follows directly from applying (16.2) to the both sides of $(ABC)^{\dagger}_{M,N} = C^{\dagger}_{Q,N}B^{\dagger}_{P,Q}A^{\dagger}_{M,P}$ and simplifying. Observe that the left-hand side of Part (b) can also be written as

$$(M^{\frac{1}{2}}ABCN^{-\frac{1}{2}})^{\dagger} = [\,(M^{\frac{1}{2}}AP^{-\frac{1}{2}})(P^{\frac{1}{2}}BQ^{-\frac{1}{2}})(Q^{\frac{1}{2}}CN^{-\frac{1}{2}})\,]^{\dagger}.$$

In that case, we see by Theorem 8.11 that Part (b) holds if and only if

$$r\begin{bmatrix} B_1B_1^*B_1 & 0 & B_1C_1 \\ 0 & -J_1J_1^*J_1 & J_1C_1^*C_1 \\ A_1B_1 & A_1A_1^*J_1 & 0 \end{bmatrix} = r(B_1) + r(J_1),$$

where

$$A_1 = M^{\frac{1}{2}}AP^{-\frac{1}{2}}, \quad B_1 = P^{\frac{1}{2}}BQ^{-\frac{1}{2}}, \quad C_1 = Q^{\frac{1}{2}}CN^{-\frac{1}{2}}, \quad J_1 = M^{\frac{1}{2}}ABCN^{-\frac{1}{2}}.$$

Simplifying this rank equality by the given condition that $M,\ N,\ P$ and Q are positive definite, we obtain the rank equality in Part (c). $\square$

Corollary 17.2. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times k}$ and $C \in \mathcal{C}^{k\times l}$ be given and let $J = ABC$. Let $P \in \mathcal{C}^{n\times n}$ and $Q \in \mathcal{C}^{k\times k}$ be two positive definite matrices. Then the following three statements are equivalent:*

(a) $(ABC)^{\dagger} = C^{\dagger}_{Q,I}B^{\dagger}_{P,Q}A^{\dagger}_{I,P}$.

(b) $(ABC)^{\dagger} = (Q^{\frac{1}{2}}C)^{\dagger}(P^{\frac{1}{2}}BQ^{-\frac{1}{2}})^{\dagger}(AP^{-\frac{1}{2}})^{\dagger}$.

(c) $r\begin{bmatrix} BQ^{-1}B^*PB & 0 & BC \\ 0 & -JJ^*J & JC^*QC \\ AB & AP^{-1}A^*J & 0 \end{bmatrix} = r(B) + r(J)$.

Proof. Follows from Theorem 17.1 by setting M and N as identity matrices. $\square$

Corollary 17.3. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times k}$, and $C \in \mathcal{C}^{k\times l}$ be given and denote $J = ABC$. Let $M \in \mathcal{C}^{m\times m}$, $N \in \mathcal{C}^{l\times l}$ be two positive definite matrices. Then the following three statements are equivalent*:

(a) $(ABC)^{\dagger}_{M,N} = C^{\dagger}_{I,N}B^{\dagger}A^{\dagger}_{M,I}$.

(b) $(M^{\frac{1}{2}}ABCN^{-\frac{1}{2}})^{\dagger} = (CN^{-\frac{1}{2}})^{\dagger}B^{\dagger}(M^{\frac{1}{2}}A)^{\dagger}$.

(c) $r\begin{bmatrix} BB^*B & 0BC & \\ 0 & -JN^{-1}J^*MJ & JN^{-1}C^*C \\ AB & AA^*MJ & 0 \end{bmatrix} = r(B) + r(J)$.

Proof. Follows from Theorem 17.1 by setting P and Q as identity matrices. $\square$

Corollary 17.4. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times k}$, and $C \in \mathcal{C}^{k\times l}$ be given with $r(A) = n$ and $r(C) = k$. Let $M \in \mathcal{C}^{m\times m}$, $N \in \mathcal{C}^{l\times l}$, $P \in \mathcal{C}^{n\times n}$, and $Q \in \mathcal{C}^{k\times k}$ be four positive definite matrices. Then the following two statements are equivalent*:

(a) $(ABC)^{\dagger}_{M,N} = C^{\dagger}_{Q,N}B^{\dagger}_{P,Q}A^{\dagger}_{M,P}$.

(b) $R(\, P^{-1}A^*MAB\,) \subseteq R(B)$ *and* $R[\,(BCN^{-1}C^*Q)^*\,] \subseteq R(B^*)$.

Proof. The given condition $r(A) = n$ and $r(C) = k$ is equivalent to $A^{\dagger}A = I_n$, $CC^{\dagger} = I_k$, and $r(ABC) = r(B)$. In that case, we can show by block elementary operations that

$$\begin{bmatrix} BQ^{-1}B^*PB & 0 & BC \\ 0 & -JN^{-1}J^*MJ & JN^{-1}C^*QC \\ AB & AP^{-1}A^*MJ & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 0 & B \\ 0 & 0 & BCN^{-1}C^*Q \\ B & P^{-1}A^*MAB & 0 \end{bmatrix}$$

are equivalent, the detailed is omitted here. This result implies that

$$\begin{aligned} & r\begin{bmatrix} BQ^{-1}B^*PB & 0 & BC \\ 0 & -JN^{-1}J^*MJ & JN^{-1}C^*QC \\ AB & AP^{-1}A^*MJ & 0 \end{bmatrix} \\ = \; & r\begin{bmatrix} 0 & 0 & B \\ 0 & 0 & BCN^{-1}C^*Q \\ B & P^{-1}A^*MAB & 0 \end{bmatrix} \\ = \; & r\begin{bmatrix} B \\ BCN^{-1}C^*Q \end{bmatrix} + r[\,B, \; P^{-1}A^*MAB\,]. \end{aligned}$$

Thus under the given condition of this corollary, Part (c) of Theorem 17.1 reduces to

$$r\begin{bmatrix} B \\ BCN^{-1}C^*Q \end{bmatrix} + r[\,B, \; P^{-1}A^*MAB\,] = 2r(B),$$

which is obviously equivalent to Part (c) of this corollary. $\square$

Corollary 17.5. *Let $A \in \mathcal{C}^{m\times m}$, $B \in \mathcal{C}^{m\times n}$, and $C \in \mathcal{C}^{n\times n}$ be given with A and C nonsingular. Let M, $P \in \mathcal{C}^{m\times m}$ and N, $Q \in \mathcal{C}^{n\times n}$ be four positive definite Hermitian matrices. Then*

(a) $(ABC)^{\dagger}_{M,N} = C^{-1}B^{\dagger}_{P,Q}A^{-1} \Leftrightarrow R(\, P^{-1}A^*MAB\,) = R(B)$ *and* $R[\,(BCN^{-1}C^*Q)^*\,] = R(B^*)$.

(b) $(ABC)^{\dagger}_{M,N} = C^{-1}B^{\dagger}A^{-1} \Leftrightarrow R(\, A^*MAB\,) = R(B)$ *and* $R[\,(BCN^{-1}C^*)^*\,] = R(B^*)$.

(c) $(ABC)^{\dagger} = C^{-1}B^{\dagger}_{P,Q}A^{-1} \Leftrightarrow R(\, P^{-1}A^*AB\,) = R(B)$ *and* $R[\,(BCC^*Q)^*\,] = R(B^*)$.

In particular, the following two identities hold

$$(ABC)^{\dagger}_{M,N} = C^{-1}B^{\dagger}_{(A^*MA),\,(CN^{-1}C^*)^{-1}}A^{-1}, \tag{17.1}$$

$$(ABC)^{\dagger} = C^{-1}B^{\dagger}_{(A^*A),\,(CC^*)^{-1}}A^{-1}. \tag{17.2}$$

Proof. Let A and C be nonsingular matrices in Corollary 17.4. We can obtain Part (a) of this corollary. Parts (a) and (b) are special cases of Part (a). The equality (17.1) follows from Part (a) by setting $P = A^*MA$ and $Q = (CN^{-1}C^*)^{-1}$. $\square$

Theorem 17.6. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times k}$, and $C \in \mathcal{C}^{k\times l}$ be given and denote $J = ABC$. Let $M \in \mathcal{C}^{m\times m}$, $N \in \mathcal{C}^{l\times l}$, $P \in \mathcal{C}^{n\times n}$, and $Q \in \mathcal{C}^{k\times k}$ be four positive definite matrices. Then the following two statements are equivalent*:

(a) $(ABC)^{\dagger}_{M,N} = (BC)^{\dagger}_{P,N}B(AB)^{\dagger}_{M,Q}$.

(b) $r\begin{bmatrix} B^* & (AB)^*MJ & B^*PBC \\ JN^{-1}(BC)^* & 0 & 0 \\ ABQ^{-1}B^* & 0 & 0 \end{bmatrix} = r(B) + r(J).$

Proof. Write ABC as $ABC = (AB)B^{\dagger}_{P,Q}(BC)$ and notice that $(B^{\dagger}_{P,Q})^{\dagger}_{Q,P} = B$. Then by Theorem 17.1, we know that

$$(ABC)^{\dagger}_{M,N} = [\,(AB)B^{\dagger}_{P,Q}(BC)\,]^{\dagger}_{M,N} = (BC)^{\dagger}_{P,N}(B^{\dagger}_{P,Q})^{\dagger}_{Q,P}(AB)^{\dagger}_{M,Q} = (BC)^{\dagger}_{P,N}B(AB)^{\dagger}_{M,Q}$$

holds if and only if

$$r\begin{bmatrix} B^{\dagger}_{P,Q}P^{-1}(B^{\dagger}_{P,Q})^*QB^{\dagger}_{P,Q} & 0 & B^{\dagger}_{P,Q}BC \\ 0 & -JN^{-1}J^*MJ & JN^{-1}(BC)^*P(BC) \\ ABB^{\dagger}_{P,Q} & ABQ^{-1}(AB)^*MJ & 0 \end{bmatrix} = r(B^{\dagger}_{P,Q}) + r(J). \tag{17.3}$$

Note by (1.5) that

$$B^{\dagger}_{P,Q}P^{-1}(B^{\dagger}_{P,Q})^*QB^{\dagger}_{P,Q} = Q^{-\frac{1}{2}}(P^{\frac{1}{2}}BQ^{-\frac{1}{2}})^{\dagger}[\,(P^{\frac{1}{2}}BQ^{-\frac{1}{2}})^{\dagger}\,]^*(P^{\frac{1}{2}}AQ^{-\frac{1}{2}})^{\dagger}P^{\frac{1}{2}}.$$

Thus by block elementary operations, we can deduce that (17.3) is equivalent to Part (c) of the theorem. The details are omitted. □

Corollary 17.7. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{n\times k}$, and $C \in \mathcal{C}^{k\times l}$ be given and denote $J = ABC$. Let $M \in \mathcal{C}^{m\times m}$, $N \in \mathcal{C}^{l\times l}$, $P \in \mathcal{C}^{n\times n}$, and $Q \in \mathcal{C}^{k\times k}$ be four positive definite matrices. If*

$$r(ABC) = r(B), \tag{17.4}$$

then the weighted Moore-Penrose inverse of the product ABC satisfies the following two equalities

$$(ABC)^{\dagger}_{M,N} = (BC)^{\dagger}_{P,N}B(AB)^{\dagger}_{M,Q}, \tag{17.5}$$

and

$$(ABC)^{\dagger}_{M,N} = (B^{\dagger}_{P,Q}BC)^{\dagger}_{Q,N}B^{\dagger}_{P,Q}(ABB^{\dagger}_{P,Q})^{\dagger}_{M,Q}. \tag{17.6}$$

Proof. Under (17.4), we know that

$$r(AB) = r(BC) = r(B),$$

which is equivalent to

$$R(BC) = R(B), \quad \text{and} \quad R(B^*A^*) = R(B^*).$$

Based on them we further obtain

$$R(B^*PBC) = R(B^*PB) = R[\,(B^*P^{\frac{1}{2}})(B^*P^{\frac{1}{2}})^*\,] = R(B^*P^{\frac{1}{2}}) = R(B^*),$$

and

$$R(BQ^{-1}B^*A^*) = R(BQ^{-1}B^*) = R[\,(BQ^{-\frac{1}{2}})(BQ^{-\frac{1}{2}})^*\,] = R(BQ^{-\frac{1}{2}}) = R(B).$$

Under these two conditions, the left-hand side of Part (b) in Theorem 17.6 reduces to $2r(B)$. Thus Part (b) in Theorem 17.6 is indentity under (17.4). Therefore we have (17.5) under (17.4). Consequently writing ABC as $ABC = (AB)B^{\dagger}_{P,Q}(BC)$ and applying (17.5) to it yields (17.6). □

Some applications of Corollary 17.7 are given below.

Corollary 17.8. *Let A, $B \in \mathcal{C}^{m\times n}$ be given, $M \in \mathcal{C}^{m\times m}$, $N \in \mathcal{C}^{n\times n}$, $P \in \mathcal{C}^{2m\times 2m}$, and $Q \in \mathcal{C}^{2n\times 2n}$ be four positive definite matrices. If A and B satisfy the rank additivity condition*

$$r(\,A + B\,) = r(A) + r(B), \tag{17.7}$$

then the weighted Moore-Penrose of $A+B$ satisfies the two equalities

$$(A+B)^{\dagger}_{M,N} = \begin{bmatrix} A \\ B \end{bmatrix}^{\dagger}_{P,N} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} [\, A, \ B \,]^{\dagger}_{M,Q}, \tag{17.8}$$

$$(A+B)^{\dagger}_{M,N} = \begin{bmatrix} A^{\dagger}_{M,N}A \\ B^{\dagger}_{M,N}B \end{bmatrix}^{\dagger}_{Q,N} \begin{bmatrix} A^{\dagger}_{M,N} & 0 \\ 0 & B^{\dagger}_{M,N} \end{bmatrix} [\, AA^{\dagger}_{M,N}, \ BB^{\dagger}_{M,N} \,]^{\dagger}_{M,P}. \tag{17.9}$$

Proof. Write $A+B$ as

$$A+B = [\, I_m, \ I_m \,] \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} I_n \\ I_n \end{bmatrix} := UDV.$$

Then the condition (17.7) is equivalent to $r(UDV) = r(D)$. Thus it turns out that

$$(UDV)^{\dagger}_{M,N} = (DV)^{\dagger}_{P,N} D (UD)^{\dagger}_{M,Q},$$

which is exactly (17.8). Next write $A+B$ as

$$A+B = [\, A, \ B \,] \begin{bmatrix} A^{\dagger}_{M,N} & 0 \\ 0 & B^{\dagger}_{M,N} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} := U_1 D_1 V_1.$$

Then the condition (17.7) is also equivalent to $r(U_1D_1V_1) = r(D_1)$. Thus it follows by (17.5) that $(U_1D_1V_1)^{\dagger}_{M,N} = (D_1V_1)^{\dagger}_{Q,N} D_1 (U_1D_1)^{\dagger}_{M,P}$, which is exactly (17.9). □

A generalization of Corollary 17.8 is presented below, the proof is omitted.

Corollary 17.9. *Let $A_1, \cdots, A_k \in \mathcal{C}^{m\times n}$ be given, and let $M \in \mathcal{C}^{m\times m}$, $N \in \mathcal{C}^{n\times n}$, $P \in \mathcal{C}^{km\times km}$, and $Q \in \mathcal{C}^{kn\times kn}$ be four positive definite Hermitian matrices. If*

$$r(\, A_1 + \cdots + A_k \,) = r(A_1) + \cdots + r(A_k), \tag{17.10}$$

then the weighted Moore-Penrose inverse of the sum satisfies the following two equalities

$$(\, A_1 + \cdots + A_k \,)^{\dagger}_{M,N} = \begin{bmatrix} A_1 \\ \vdots \\ A_k \end{bmatrix}^{\dagger}_{P,N} \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix} [\, A_1, \ \cdots, \ A_k \,]^{\dagger}_{M,Q}, \tag{17.11}$$

$$\left(\sum_{t=1}^{k} A_t \right)^{\dagger}_{M,N} = \begin{bmatrix} (A_1)^{\dagger}_{M,N}A_1 \\ \vdots \\ (A_k)^{\dagger}_{M,N}A_k \end{bmatrix}^{\dagger}_{Q,N} \begin{bmatrix} (A_1)^{\dagger}_{M,N} & & \\ & \ddots & \\ & & (A_k)^{\dagger}_{M,N} \end{bmatrix} [\, A_1(A_1)^{\dagger}_{M,N}, \ \cdots, \ A_k(A_k)^{\dagger}_{M,N} \,]^{\dagger}_{M,P}. \tag{17.12}$$

Corollary 17.10. *Let $A \in \mathcal{C}^{m\times n}$, $B \in \mathcal{C}^{m\times k}$, $C \in \mathcal{C}^{l\times n}$, $A \in \mathcal{C}^{l\times k}$ be given, $M, \ P \in \mathcal{C}^{(m+l)\times(m+l)}$, $N, \ Q \in \mathcal{C}^{(n+k)\times(n+k)}$ be four positive definite matrices. If*

$$r \begin{bmatrix} A & B \\ C & D \end{bmatrix} = r(A), \tag{17.13}$$

or equivalently $AA^{\dagger}B = B$, $CA^{\dagger}A = C$ and $D = CA^{\dagger}B$, then

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger}_{M,N} = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}^{\dagger}_{P,N} \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix}^{\dagger}_{M,Q}. \tag{17.14}$$

In particular,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger}_{M,N} = [\, A, \ B\,]^{\dagger}_{I,N} A \begin{bmatrix} A \\ C \end{bmatrix}^{\dagger}_{M,I}. \tag{17.15}$$

Proof. Under (17.13), we see that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ CA^{\dagger} & I_l \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_n & A^{\dagger}B \\ 0 & I_k \end{bmatrix} := ULV.$$

Thus by Corollary 17.6, we obtain

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\dagger}_{M,N} = (LV)^{\dagger}_{P,N} L (UL)^{\dagger}_{M,Q},$$

which is exactly (17.14). When $P = I_{m+l}$ and $Q = I_{n+k}$, we have

$$\begin{aligned} \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}^{\dagger}_{I,N} \begin{bmatrix} I_m \\ 0 \end{bmatrix} &= N^{-\frac{1}{2}} \left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} N^{-\frac{1}{2}} \right)^{\dagger} \begin{bmatrix} I_m \\ 0 \end{bmatrix} \\ &= N^{-\frac{1}{2}} \begin{bmatrix} (\,[\, A, \ B\,]\, N^{-\frac{1}{2}}\,)^{\dagger} \\ [\,0, \ 0\,]^{\dagger} \end{bmatrix} \begin{bmatrix} I_m \\ 0 \end{bmatrix} \\ &= N^{-\frac{1}{2}} ([\, A, \ B\,] N^{-\frac{1}{2}})^{\dagger} = [\, A, \ B\,]^{\dagger}_{I,N}. \end{aligned}$$

Similarly we can deduce

$$[\, I_n, \ 0\,] \begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix}^{\dagger}_{M,I} = \begin{bmatrix} A \\ C \end{bmatrix}^{\dagger}_{M,I}.$$

Putting both of them in (17.14) yields (17.15). $\square$

Chapter 18

Extreme ranks of $A - BXC$

The basic tool for establishing the whole work in the monograph is the rank formula (2.1) for the Schur complement $D - CA^{\dagger}B$. Motivated by (2.1), one might naturally consider the rank of a generalized Schur complement $D - CA^{-}B$, where A^{-} is an inner inverse of A. Since A^{-} is not unique in general, the rank of $D - CA^{-}B$ will depend on the choice of A^{-}. Thus a fundamental problem related to a generalized Schur complement $D - CA^{-}B$ is to find its maximal and minimal possible ranks with respect to the choice of A^{-}. Notice that the general expression of A^{-} is $A^{-} = A^{\dagger} + F_AV + WE_A$, where both V and W are arbitrary matrices. As a consequence,

$$D - CA^{-}B = D - CA^{\dagger}B - CF_AVB - CWE_AB.$$

This expression implies that $D - CA^{-}B$ is in fact a matrix expression with two independent variant matrices. This fact motivates us to consider another basic problem in matrix theory—maximal and minimal possible ranks of linear matrix expressions with variant matrices. In this chapter, we consider the simplest case— the maximal and the minimal ranks of the matrix expression $A - BXC$ with respect to the variant matrix X and then discuss some related topics. Throughout the symbols E_A and F_A stand for the two oblique projectors $E_A = I - AA^{-}$ and $F_A = I - A^{-}A$ induced by A.

The following result is well known (see, e.g., [118]).

Lemma 18.1. *Suppose $BXC = A$ is a linear matrix equation over an arbitrary field $\mathcal{F}$, where $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$, $C \in \mathcal{F}^{l\times n}$ are given. Then it is consistent if and only if $R(A) \subseteq R(B)$ and $R(A^T) \subseteq R(C^T)$, or equivalently $BB^{-}AC^{-}C = A$. In that case, the general solution of $BXC = A$ can be expressed as*

$$X = B^{-}AC^{-} + U - B^{-}BUCC^{-}, \quad or \quad X = B^{-}AC^{-} + F_BV + WE_A,$$

where U, V and W are arbitrary matrices. In particular, the solution of $BXC = A$ is unique if and only if B has full column rank and C has full row rank.

In order to determine the maximal and minimal ranks of $A - BXC$ with respect to X, we first establish two rank identities for $A - BXC$ through (1.4) and (1.5).

Theorem 18.2. *The matrix expression $A - BXC$ satisfies the rank identity*

$$r(A - BXC) = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,] - r(M) + r[\,E_{T_1}(\,X + TM^{-}S\,)F_{S_1}\,], \tag{18.1}$$

where $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$, $T = [\,0,\ I_k\,]$ *and* $S = \begin{bmatrix} 0 \\ I_l \end{bmatrix}$, $T_1 = TF_M$, *and* $S_1 = E_MS$.

Proof. It is easy to verify by block elementary operations of matrix that

$$r(A - BXC) = r\begin{bmatrix} A & B & 0 \\ C & 0 & I_l \\ 0 & I_k & -X \end{bmatrix} - k - l = r\begin{bmatrix} M & S \\ T & -X \end{bmatrix} - k - l. \tag{18.2}$$

Applying (1.6) to the block matrix in it, we find that

$$\begin{aligned} r\begin{bmatrix} M & S \\ T & -X \end{bmatrix} &= r\begin{bmatrix} M \\ T \end{bmatrix} + r[\, M,\ S\,] - r(M) + r[\, E_{T_1}(\, X + TM^{-}S\,)F_{S_1}\,] \\ &= r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A,\ B\,] + k + l - r(M) + r[\, E_{T_1}(\, X + TM^{-}S\,)F_{S_1}\,]. \end{aligned}$$

Putting it in (18.2) yields (18.1). $\square$.

Theorem 18.3. *The matrix expression $A - BXC$ satisfies the rank identity*

$$r(A - BXC) = r[\, A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} + r(E_{A_2}AF_{A_1} - E_{A_2}BXCF_{A_1}), \tag{18.3}$$

where $A_1 = E_BA,\ A_2 = AF_C$, and the matrix equation $E_{A_2}BXCF_{A_1} = E_{A_2}AF_{A_1}$ is consistent.

Proof. We first establish the following rank equality

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A,\ B\,] - r(A) + r(E_{A_2}AF_{A_1}). \tag{18.4}$$

Observe that

$$r(E_BAF_C) = r\begin{bmatrix} A & AF_C \\ E_BA & 0 \end{bmatrix} - r(A),$$

and also observe from (1.4) that

$$r\begin{bmatrix} A & AF_C \\ E_BA & 0 \end{bmatrix} = r(E_BA) + r(AF_C) + r(E_{A_2}AF_{A_1}).$$

We obtain $r(E_BAF_C) = r(E_BA) + r(AF_C) - r(A) + r(E_{A_2}AF_{A_1})$. Putting it in (1.4) and applying (1.2) and (1.3), we get (18.4). Next replace the matrix A in (18.4) by $p(X) = A - BXC$ and notice that

$$r\begin{bmatrix} A - BXC & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}, \qquad r\begin{bmatrix} A - BXC \\ C \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix},$$
$$r[\, A - BXC,\ B\,] = r[\, A,\ B\,], \qquad E_B(A - BXC) = E_BA, \qquad (A - BXC)F_C = AF_C.$$

Then (18.4) becomes

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A,\ B\,] - r(A - BXC) + r(E_{A_2}AF_{A_1} - E_{A_2}BXCF_{A_1}),$$

establishing (18.3). On the other hand, from $E_{A_2}A_2 = 0$ and $A_1F_{A_1} = 0$ we can deduce that $E_{A_2}AC^{-}C = E_{A_2}A$ and $BB^{-}AF_{A_1} = AF_{A_1}$. Thus $R(E_{A_2}AF_{A_1}) = R(E_{A_2}BB^{-}AF_{A_1}) \subseteq R(E_{A_2}B)$ and $R[\,(E_{A_2}AF_{A_1})^T\,] = R[\,(E_{A_2}AC^{-}CF_{A_1})^T] \subseteq R[\,(CF_{A_1})^T\,]$. Both of them imply that the matrix equation $E_{A_2}BXCF_{A_1} = E_{A_2}AF_{A_1}$ is consistent. $\square$

On the basis of (18.3), we establish the main result of the chapter.

Theorem 18.4. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ and $C \in \mathcal{F}^{l\times n}$ be given. Then*

(a) *The maximal rank of $A - BXC$ with respect to X is*

$$\max_X r(\, A - BXC\,) = \min\left\{ r[\, A,\ B\,], \quad r\begin{bmatrix} A \\ C \end{bmatrix} \right\}. \tag{18.5}$$

(b) *The minimal rank of $A - BXC$ with respect to X is*

$$\min_X r(\, A - BXC\,) = r[\, A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}. \tag{18.6}$$

(c) *The general expression of X satisfying* (18.5) *can be written as*

$$X = -TM^{-}S + U, \tag{18.7}$$

where U is chosen such that $r(E_{T_1}UF_{S_1}) = \min\{\, r(E_{T_1}),\ r(F_{S_1})\,\}$, where $M,\ S,\ T,\ S_1$ and T_1 are defined in (18.1).

(d) *The matrix X satisfying* (18.6) *is determined by the matrix equation $E_{T_1}(\, X + TM^{-}S\,)F_{S_1} = 0$, and can be written as*

$$X = -TM^{-}S + T_1V + WS_1, \tag{18.8}$$

where V and W are arbitrary.

Proof. Eq. (18.2) implies that

$$\max_X r(\, A - BXC\,) = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A,\ B\,] - r(M) + \max_X r[\, E_{T_1}(\, X + TM^{-}S\,)F_{S_1}\,], \tag{18.9}$$

$$\min_X r(\, A - BXC\,) = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A,\ B\,] - r(M) + \min_X r[\, E_{T_1}(\, X + TM^{-}S\,)F_{S_1}\,]. \tag{18.10}$$

It is obvious that

$$\max_X r[\, E_{T_1}(\, X + TM^{-}S\,)F_{S_1}\,] = \max_Y r(\, E_{T_1}YF_{S_1}\,) = \min\{\ r(E_{T_1}), \quad r(F_{S_1})\ \}, \tag{18.11}$$

and the matrix satisfying it can be written as (18.7). According to (1.2) and (1.3), we find that

$$r(E_{T_1}) = k - r(T_1) = k - r(TF_M) = k - r\begin{bmatrix} M \\ T \end{bmatrix} + r(M) = r(M) - r\begin{bmatrix} A \\ C \end{bmatrix},$$

$$r(F_{S_1}) = l - r(S_1) = k - r(E_MS) = k - r[\, M,\ S\,] + r(M) = r(M) - r[\, A,\ B\,].$$

Putting both of them in (18.11) and then (18.11) in (18.9) yields (18.5). The results in (18.6) and (18.8) are direct consequences of (18.10). □.

The maximal and the minimal ranks of $A - BXC$ with respect to X can also be determined through the rank identity (18.3). In that case, the matrix X satisfying (18.5) and (18.6) can respectively be determined by the expression matrix $E_{A_2}AF_{A_1} - E_{A_2}BXCF_{A_1}$, where the corresponding matrix equation $E_{A_2}BXCF_{A_1} = E_{A_2}AF_{A_1}$ is consistent.

Corollary 18.5. *Let $p(X) = A - BXC$ be given over $\mathcal{F}$ with $B \neq 0$ and $C \neq 0$. Then*

(a) *The rank of $A - BXC$ is invariant with respect to the choice of X if and only if*

$$R\begin{bmatrix} B \\ 0 \end{bmatrix} \subseteq R\begin{bmatrix} A \\ C \end{bmatrix} \quad \text{or} \quad R([\, C,\, 0\,]^T) \subseteq R([\, A,\ B\,]^T). \tag{18.12}$$

(b) *The range $R(\, A - BXC\,)$ is invariant with respect to the choice of X if and only if*

$$R\begin{bmatrix} B \\ 0 \end{bmatrix} \subseteq R\begin{bmatrix} A \\ C \end{bmatrix}. \tag{18.13}$$

(c) *The range $R[\,(\, A - BXC\,)^T\,]$ is invariant with respect to the choice of X if and only if*

$$R(\,[\, C,\ 0\,]^T\,) \subseteq R(\,[\, A,\ B\,]^T\,). \tag{18.14}$$

(d) *The rank of $A - BXC$ is invariant with respect to the choice of X if and only if the range $R(\, A - BXC\,)$ is invariant with respect to the choice of X or the range $R[\,(\, A - BXC\,)^T\,]$ is invariant with respect to the choice of X.*

Proof. From (18.5) and (18.6), we obtain

$$\max_X r(\, A - BXC\,) - \min_X r(\, A - BXC\,) = \min\left\{ r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r\begin{bmatrix} A \\ C \end{bmatrix},\ \ r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r[\, A,\ B\,] \right\}.$$

Let its right-hand side be zero, then we get (18.12). To show Part (b), we use a fundamental fact that two matrices P and Q have the same range, i.e., $R(P) = R(Q)$, if and only if $r[\,P,\ Q\,] = r(P) = r(Q)$. Applying this fact to $A - BXC$, we know that the range $R(A - BXC)$ is invariant with respect to the choice of X if and only if

$$r[\,A - BXC,\ A - BYC\,] = r(\,A - BXC\,) = r(\,A - BYC\,) \tag{18.15}$$

holds for all X and Y. Obviously this equality holds for all X and Y if and only if

$$r(\,A - BXC\,) = r(A), \tag{18.16}$$

holds for all X, and

$$r[\,A - BXC,\ A - BYC\,] = r\left([\,A,\ A\,] - B[\,X,\ Y\,]\begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}\right) = r(A) \tag{18.17}$$

holds for all X and Y. According to Part (a), the equality (18.16) holds for all X if and only if (18.12) holds, and the equality (18.17) holds if and only if

$$r\begin{bmatrix} A & A & B \\ C & 0 & 0 \\ 0 & C & 0 \end{bmatrix} = r\begin{bmatrix} A & A \\ C & 0 \\ 0 & C \end{bmatrix} \quad \text{or} \quad r\begin{bmatrix} A & A & B \\ C & 0 & 0 \\ 0 & C & 0 \end{bmatrix} = r[\,A,\ A,\ B\,],$$

that is,

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} \quad \text{or} \quad r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r[\,A,\ B\,] - r(C). \tag{18.18}$$

Note that $B \neq 0$ and $C \neq 0$. Thus combining (18.12) with (18.18), we know that (18.15) holds if and only if

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix},$$

which is equivalent to (18.13). Similarly we can show Part (c). Contrasting Parts (a)—(c) yields Part (b). □

Corollary 18.6. *The matrix satisfying* (18.6) *is unique if and only if*

$$r(B) = k, \quad r(C) = l, \quad \textit{and} \quad r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r(B) = r[\,A,\ B\,] + r(C). \tag{18.19}$$

In that case, the unique matrix satisfying (18.6) *is*

$$X = -[\,0,\ I_k\,]\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^{-}\begin{bmatrix} 0 \\ I_l \end{bmatrix}. \tag{18.20}$$

Proof. The matrix satisfying (18.6) is unique if and only if the solution to the equation $E_{T_1}(\,X + TM^{-}S\,)F_{S_1} = 0$ is unique, which is equivalent to

$$r(E_{T_1}) = k \quad \text{and} \quad r(F_{S_1}) = l. \tag{18.21}$$

Recall that

$$r(E_{T_1}) = r(M) - r\begin{bmatrix} A \\ C \end{bmatrix}, \quad \text{and} \quad r(F_{S_1}) = r(M) - r[\,A,\ B\,],$$

and $r(B) \leq k$ and $r(C) \leq l$. Thus (18.21) is equivalent to (18.19), and the unique matrix is (18.20). □.

Corollary 18.7. *The following four statements are equivalent*:

(a) $\min_X r(A - BXC) = r(A)$.

(b) $r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,] - r(A)$.

(c) $E_{T_1}TM^{-}SF_{S_1} = 0$.

(d) $E_{A_2}AF_{A_1} = 0$.
(e) $E_{C_1}CA^-BF_{B_1} = 0$, *where* $A_1 = E_BA,\ A_2 = AF_C,\ B_1 = E_AB,\ C_1 = CF_A$.

Proof. Follows immediately from the combination of (18.6), (18.1), (18.2) and (1.6). □

In the remainder of this section, we present some equivalent statements for the results in Theorem 18.4. Suppose that $B \in \mathcal{F}^{m\times k}$, $C \in \mathcal{F}^{l\times n}$, $P \in \mathcal{F}^{s\times m}$, $Q \in \mathcal{F}^{n\times t}$, and let Θ be the matrix set

$$\Theta = \{\, Z \in \mathcal{F}^{m\times n} \mid R(Z) \subseteq R(B) \ \text{ and } \ R(Z^T) \subseteq R(C^T)\,\}. \tag{18.22}$$

Then we have the following results.

Theorem 18.8. *Let $A \in \mathcal{F}^{m\times n}$ be given and Θ be defined in* (18.21). *Then*
(a) *The maximal rank of $A - Z$ subject to $Z \in \Theta$ is*

$$\max_{Z\in\Theta} r(A - Z) = \min\left\{ r[\,A,\ B\,], \qquad r\begin{bmatrix} A \\ C \end{bmatrix} \right\}, \tag{18.23}$$

and the matrix Z satisfying (18.23) *can be written in the form*

$$Z = -[\,0,\ B\,]\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^- \begin{bmatrix} 0 \\ C \end{bmatrix} - BUC, \tag{18.24}$$

where U is chosen such that $r(E_{T_1}UF_{S_1}) = \min\{\, r(E_{T_1}), \quad r(F_{S_1})\,\}$, where $M,\ S,\ T,\ S_1$ and T_1 are as in (18.1).
(b) *The minimal rank of $A - Z$ subject to $Z \in \Theta$ is*

$$\min_{Z\in\Theta} r(A - Z) = r[\,A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}, \tag{18.25}$$

and the general expression of the matrix Z satisfying (18.24) *can be written as*

$$Z = -[\,0,\ B\,]\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^- \begin{bmatrix} 0 \\ C \end{bmatrix} + BT_1VC + BWS_1C, \tag{18.26}$$

where V and W are arbitrary.

Proof. From the structure of Θ in (18.22) we easily see that Θ can equivalently be expressed as

$$\Theta = \{\, Z = BXC \mid X \in \mathcal{F}^{k\times l}\,\}.$$

Thus the rank of $A - Z$ subject to $Z \in \Theta$ can be written as

$$r(A - Z) = r(A - BXC) = r[\,A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} + r[\,E_{T_1}(\,X + TM^-S\,)F_{S_1}\,]. \tag{18.27}$$

In this case, applying Theorem 18.4 to this equality, we obtain the desired results in the theorem. □

The matrix $Z \in \Theta$ satisfying (18.25) is well known as a *shorted matrix* of A relative to Θ. Thus (18.26) is in fact the general expression of shorted matrices of A relative to Θ. One of the most important aspects on shorted matrices is concerning their uniqueness, which has been well examined by lots of authors (see, e.g., [2], [23], [102], [106]). Now from the general result in Theorem 18.6 and 18.8(b) and we easily get the following known result.

Theorem 18.9[102]. *Let $A \in \mathcal{F}^{m\times n}$ be given and Θ be defined in* (18.22). *Then the shorted matrix of A relative to Θ is unique if and only if*

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r(B) = r[\,A,\ B\,] + r(C). \tag{18.28}$$

In that case, the unique shorted matrix is

$$Z = -[\,0,\ B\,]\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^- \begin{bmatrix} 0 \\ C \end{bmatrix}, \tag{18.29}$$

and this matrix is invariant with respect to the choice of the inner inverse in it.

Chapter 19

Extreme ranks of $A - B_1X_1C_1 - B_2X_2C_2$

In order to find the maximal and the minimal ranks of $D - CA^-B$ with respect to A^-, we need to know maximal and minimal ranks of

$$p(X_1,\ X_2) = A - B_1X_1C_1 - B_2X_2C_2 \tag{19.1}$$

under the two conditions

$$R(B_1) \subseteq R(B_2) \ \text{ and } \ R(C_2^T) \subseteq R(C_1^T), \tag{19.2}$$

where A, B_1, B_2, C_1 and C_2 are given, X_1 and X_2 are two independent variant matrices over $\mathcal{F}$.

Theorem 19.1. *Let $p(X_1,\ X_2)$ be given by* (19.1) *and* (19.2). *Then the maximal rank of $p(X_1,\ X_2)$ with respect to X_1 and X_2 is*

$$\max_{X_1,\,X_2} r[\,p(X_1,\ X_2)\,] = \min\left\{ r[\,A,\ B_2\,],\quad r\begin{bmatrix} A \\ C_1 \end{bmatrix},\quad r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} \right\}. \tag{19.3}$$

Proof. Applying (18.5) to $p(X_1,\ X_2)$ in (19.1) we first obtain

$$\begin{aligned} \max_{X_2} r[\,p(X_1,\ X_2)\,] &= \min\left\{ r(\,A - B_1X_1C_1,\ B_2\,),\ \ r\begin{bmatrix} A - B_1X_1C_1 \\ C_2 \end{bmatrix} \right\} \\ &= \min\left\{ r[\,A,\ B_2\,],\quad r\begin{bmatrix} A - B_1X_1C_1 \\ C_2 \end{bmatrix} \right\}. \end{aligned}$$

Next applying (18.5) to $\begin{bmatrix} A - B_1X_1C_1 \\ C_2 \end{bmatrix}$, we obtain

$$\max_{X_1} r\begin{bmatrix} A - B_1X_1C_1 \\ C_2 \end{bmatrix} = \max_{X_1} r\left(\begin{bmatrix} A \\ C_2 \end{bmatrix} - \begin{bmatrix} B_1 \\ 0 \end{bmatrix} X_1C_1 \right) = \min\left\{ r\begin{bmatrix} A \\ C_1 \end{bmatrix},\quad r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} \right\}.$$

Combining the above two results yields (19.3). $\square$

Theorem 19.2. *Let $p(X_1,\ X_2)$ be given by* (19.1) *and* (19.2). *Then the minimal rank of $p(X_1,\ X_2)$ with respect to X_1 and X_2 is*

$$\min_{X_1,\,X_2} r[\,p(X_1,\ X_2)\,] = r[\,A,\ B_2\,] + r\begin{bmatrix} A \\ C_1 \end{bmatrix} + r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix}. \tag{19.4}$$

Proof. Applying (18.6) to $p(X_1,\ X_2)$ in (19.1) we first obtain

$$\begin{aligned} \min_{X_2} r[\,p(X_1,\ X_2)\,] &= r[\,A - B_1X_1C_1,\ B_2\,] + r\begin{bmatrix} A - B_1X_1C_1 \\ C_2 \end{bmatrix} - r\begin{bmatrix} A - B_1X_1C_1 & B_2 \\ C_2 & 0 \end{bmatrix} \\ &= r[\,A,\ B_2\,] + r\begin{bmatrix} A - B_1X_1C_1 \\ C_2 \end{bmatrix} - r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix}. \end{aligned}$$

Next applying (18.6) to $\begin{bmatrix} A - B_1X_1C_1 \\ C_2 \end{bmatrix}$, we find

$$\min_{X_1} r\begin{bmatrix} A - B_1X_1C_1 \\ C_2 \end{bmatrix} = \min_{X_1} r\left(\begin{bmatrix} A \\ C_2 \end{bmatrix} - \begin{bmatrix} B_1 \\ 0 \end{bmatrix} X_1C_1\right) = r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} + r\begin{bmatrix} A \\ C_1 \end{bmatrix} - r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix}.$$

Combining the above two results yields (19.4). □

The matrices X_1 and X_2 satisfying (19.3) and (19.4) can also be derived through the two expressions in (18.7) and (18.8). But their expressions are somewhat complicated in form and are omitted them here.

Eq. (19.4) can also be written as

$$\begin{aligned} \min_{X_1, X_2} r[\,p(X_1\, X_2)\,] \;&=\; \left(r[\,A,\ B_2\,] + r\begin{bmatrix} A \\ C_1 \end{bmatrix} - r\begin{bmatrix} A & B_2 \\ C_1 & 0 \end{bmatrix}\right) \\ &\quad + \left(r\begin{bmatrix} A & B_2 \\ C_1 & 0 \end{bmatrix} + r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix}\right). \end{aligned}$$

It is easy to verify that under (19.2) the two quantities in the parentheses on the right hand-side of the above equality are nonnegative. Thus the right hand-side of (19.4) is also nonnegative, although this is not evident from its expression.

Some direct consequences of Theorems 19.1 and 19.2 are given below.

Corollary 19.3. *Let $p(X_1,\, X_2)$ be given by* (19.1) *and* (19.2). *Then the rank of $p(X_1,\, X_2)$ is invariant with respect to the choice of X_1 and X_2 if and only if*

$$r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C_1 \end{bmatrix} \quad \text{and} \quad r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} = r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix}, \tag{19.5}$$

or

$$r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} = r[\,A,\ B_2\,] \quad \text{and} \quad r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} = r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix}, \tag{19.6}$$

or

$$r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C_1 \end{bmatrix} \quad \text{and} \quad r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} = r[\,A,\ B_2\,]. \tag{19.7}$$

Proof. Combining (19.3) and (19.4), we obtain the following

$$\max_{X_1, X_2} r[\,p(X_1,\, X_2)\,] - \min_{X_1, X_2} r[\,p(X_1,\, X_2)\,] = \min\{\, s_1, \quad s_2, \quad s_3\,\},$$

where

$$\begin{aligned} s_1 &= r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} + r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} - r\begin{bmatrix} A \\ C_1 \end{bmatrix} - r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix}, \\ s_2 &= r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} + r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} - r[\,A,\ B_2\,] - r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix}, \\ s_3 &= r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} + r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} - r\begin{bmatrix} A \\ C_1 \end{bmatrix} - r[\,A,\ B_2\,]. \end{aligned}$$

Let the right-hand side of the above equality be zero. Then we obtain (19.5)—(19.7). □

Corollary 19.4. *Let $p(X_1,\, X_2)$ be given by* (19.1) *and* (19.2) *with $B_1 \neq 0$ and $C_2 \neq 0$. Then*

(a) *The range $R[\,p(X_1,\, X_2)\,]$ is invariant with respect to the choice of X_1 and X_2 if and only if*

$$r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C_1 \end{bmatrix} \quad \text{and} \quad r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} = r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix}. \tag{19.8}$$

(b) *The range $R[\,p^T(X_1,\, X_2)\,]$ is invariant with respect to the choice of X_1 and X_2 if and only if*

$$r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} = r[\,A,\ B_2\,] \quad \text{and} \quad r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} = r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix}. \tag{19.9}$$

Proof. It is obvious that the range $R[\,p(X_1,\ X_2)\,]$ is invariant with respect to the choice of X_1 and X_2 if and only if

$$r[\,p(X_1,\ X_2),\ p(Y_1,\ Y_2)\,] = r[\,p(X_1,\ X_2)\,] = r[\,p(Y_1,\ Y_2)\,] = r(A) \tag{19.10}$$

holds for all $X_1,\ X_2,\ Y_1$ and Y_2. By Corollary 19.3, $r[\,p(X_1,\ X_2)\,] = r(A)$ holds for all $X_1,\ X_2$ if and only if one of (19.5)—(19.7) holds. On the other hand,

$$[\,p(X_1,\ X_2),\ p(Y_1,\ Y_2)\,] = [\,A,\ A\,] - B_1[\,X_1,\ Y_1\,]\begin{bmatrix} C_1 & 0 \\ 0 & C_1 \end{bmatrix} - B_2[\,X_2,\ Y_2\,]\begin{bmatrix} C_2 & 0 \\ 0 & C_2 \end{bmatrix}.$$

Then according to Corollary 19.3, this expression satisfies (19.10) if and only if

$$r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C_1 \end{bmatrix} \quad \text{and} \quad r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} = r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix}, \tag{19.11}$$

or

$$r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} + r(C_2) = r[\,A,\ B_2\,] \quad \text{and} \quad r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} + r(C_2) = r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix}, \tag{19.12}$$

or

$$r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C_1 \end{bmatrix} \quad \text{and} \quad r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} + r(C_2) = r[\,A,\ B_2\,]. \tag{19.13}$$

Contrasting (19.11)—(19.13) with (19.5)—(19.7) and noticing the condition $B_1 \neq 0$ and $C_2 \neq 0$, we find that (19.10) holds if and only if (19.11), i.e., (19.8) holds. Similarly we can show Part (b). □

If one of B_1, B_2, C_1 and C_2 in (19.1) is a null matrix, then $p(X_1,\ X_2)$ becomes an expression with a single variant matrix in it. In that case, the range invariance criterion is listed in Corollary 18.5(b) and (c).

Corollary 19.5. *Let* $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ *and* $C \in \mathcal{F}^{l\times n}$ *be given. Then*

$$\max_{X,\,Y} r(\,A - BX - YC\,) = \min\left\{ m, \quad n, \quad r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} \right\}, \tag{19.14}$$

$$\min_{X,\,Y} r(\,A - BX - YC\,) = r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r(B) - r(C). \tag{19.15}$$

A pair of matrices X and Y satisfying (19.5) *can be written as*

$$X = B^-A + UC + (\,I_k - B^-B\,)V_1, \quad Y = (\,I_m - BB^-\,)AC^- - BU + V_2(\,I_l - CC^-\,), \tag{19.16}$$

where U, V_1 and V_2 are arbitrary.

Proof. Eqs. (19.14) and (19.15) follow immediately from (19.3) and (19.4). Putting (19.16) in $A - BX - YC$ yields

$$A - BX - YC = (\,I_m - BB^-\,)A(\,I_n - C^-C\,).$$

Thus we have (19.15) by (1.4). □

Chapter 20

Extreme ranks of $A_1 - B_1XC_1$ subject to $B_2XC_2 = A_2$

Based on the results in Chapter 19, we are now able to find the maximal and the minimal ranks of $A_1 - B_1XC_1$ subject to a consistent linear matrix equation $B_2XC_2 = A_2$. The corresponding results will widely be used in the sequel.

Theorem 20.1. *Suppose that the matrix equation $B_2XC_2 = A_2$ is a consistent. Then*

(a) *The maximal rank of $p(X) = A_1 - B_1XC_1$ subject to $B_2XC_2 = A_2$ is*

$$\max_{B_2XC_2=A_2} r[\,p(X)\,] = \min\left\{ r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} - r(B_2) - r(C_2), \quad r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix}, \quad r[\,A_1, \ B_1\,] \right\}. \tag{20.1}$$

(b) *The minimal rank of $p(X) = A_1 - B_1XC_1$ subject to $B_2XC_2 = A_2$ is*

$$\min_{B_2XC_2=A_2} r[\,p(X)\,]$$
$$= r[\,A_1, \ B_1\,] + r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix} - r\begin{bmatrix} A_1 & B_1 & 0 \\ C_1 & 0 & C_2 \end{bmatrix} - r\begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \\ 0 & B_2 \end{bmatrix} + r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix}. \tag{20.2}$$

Proof. Note from Lemma 18.1 that the general solution of the consistent linear matrix equation $B_2XC_2 = A_2$ can be written as $X = X_0 + F_{B_2}V + WE_{C_2}$, where $X_0 = B_2^- A_2 C_2^-$, V and W are arbitrary. Putting it in $p(X) = A_1 - C_1XB_1$ yields

$$p(X) = A - B_1F_{B_2}VC_1 - B_1WE_{C_2}C_1,$$

where $A = A_1 - B_1X_0C_1$. Observe that $R(B_1F_{B_2}) \subseteq R(B_1)$ and $R[(E_{C_2}C_1)^T] \subseteq R[(C_1)^T]$. Thus it follows by (19.3) and (19.4) that

$$\max_{B_2XC_2=A_2} r[\,p(X)\,]$$
$$= \max_{V,\,W} r(\,A - B_1F_{B_2}VC_1 - B_1WE_{C_2}C_1\,)$$
$$= \min\left\{ r[\,A, \ B_1\,], \quad r\begin{bmatrix} A \\ C_1 \end{bmatrix}, \quad r\begin{bmatrix} A & B_1F_{B_2} \\ E_{C_2}C_1 & 0 \end{bmatrix} \right\},$$

and

$$\min_{B_2XC_2=A_2} r[\,p(X)\,]$$
$$= \min_{V,\,W} r(\,A - B_1F_{B_2}VC_1 - B_1WE_{C_2}C_1\,)$$
$$= r[\,A, \ B_1\,] + r\begin{bmatrix} A \\ C_1 \end{bmatrix} + r\begin{bmatrix} A & B_1F_{B_2} \\ E_{C_2}C_1 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1F_{B_2} \\ C_1 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 \\ E_{C_2}C_1 & 0 \end{bmatrix}.$$

Simplifying the ranks of the block matrices by Lemma 1.1, we see that

$$r[\,A,\ B_1\,] = r[\,A_1 - B_1X_0C_1,\ B_1\,] = r[\,A_1,\ B_1\,], \quad r\begin{bmatrix} A \\ C_1 \end{bmatrix} = r\begin{bmatrix} A_1 - B_1X_0C_1 \\ C_1 \end{bmatrix} = r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix},$$

$$\begin{aligned} r\begin{bmatrix} A & B_1F_{B_2} \\ E_{C_2}C_1 & 0 \end{bmatrix} &= r\begin{bmatrix} A_1 - B_1X_0C_1 & B_1 & 0 \\ C_1 & 0 & C_2 \\ 0 & B_2 & 0 \end{bmatrix} - r(B_2) - r(C_2) \\ &= r\begin{bmatrix} A_1 & B_1 & 0 \\ C_1 & 0 & C_2 \\ 0 & B_2 & -A_2 \end{bmatrix} - r(B_2) - r(C_2), \end{aligned}$$

$$r\begin{bmatrix} A & B_1F_{B_2} \\ C_1 & 0 \end{bmatrix} = r\begin{bmatrix} A_1 - B_1X_0C_1 & B_1 \\ C_1 & 0 \\ 0 & B_2 \end{bmatrix} - r(B_2) = r\begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \\ 0 & B_2 \end{bmatrix} - r(B_2),$$

$$r\begin{bmatrix} A & B_1 \\ E_{C_2}C_1 & 0 \end{bmatrix} = r\begin{bmatrix} A_1 - B_1X_0C_1 & B_1 & 0 \\ C_1 & 0 & C_2 \end{bmatrix} - r(C_2) = r\begin{bmatrix} A_1 & B_1 & 0 \\ C_1 & 0 & C_2 \end{bmatrix} - r(C_2).$$

Putting them in the above two rank equalities yields (20.1) and (20.2). □

Eq. (20.2) can also be written as

$$\begin{aligned} \min_{B_2XC_2=A_2} r(\,A_1 - B_1XC_1\,) &= \left(r[\,A_1,\ B_1\,] + r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix} - r\begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \end{bmatrix} \right) \\ &+ \left(r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} + r\begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \end{bmatrix} - r\begin{bmatrix} A_1 & B_1 & 0 \\ C_1 & 0 & C_2 \end{bmatrix} - r\begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \\ 0 & B_2 \end{bmatrix} \right), \end{aligned}$$

and the two quantities in the parentheses on the right hand-side of the above equality are nonnegative.

Some direct consequences are given below.

Corollary 20.2. *Suppose that $B_1XC_1 = A_1$ and $B_2XC_2 = A_2$ are consistent, respectively. Then*

$$\max_{B_2XC_2=A_2} r(\,A - B_1XC_1\,) = \min\left\{ r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} - r(B_2) - r(C_2), \quad r(C_1), \quad r(B_1) \right\}, \tag{20.3}$$

and

$$\min_{B_2XC_2=A_2} r(\,A - B_1XC_1\,) = r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} - r\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} - r[\,C_1,\ C_2\,]. \tag{20.4}$$

Proof. The consistency of $B_1XC_1 = A_1$ implies that $R(A_1) \subseteq R(B_1)$ and $R(A_1^T) \subseteq R(C_1^T)$, the consistency of $B_2XC_2 = A_2$ implies that $R(A_2) \subseteq R(B_2)$, and $R(A_2^T) \subseteq R(C_2^T)$. In that case, (20.1) and (20.2) simplify to (20.3) and (20.4). □

Notice a simple fact that the pair of matrix equations $B_1XC_1 = A_1$ and $B_2XC_2 = A_2$ have a common solution if and only if $B_1XC_1 = A_1$ and $B_2XC_2 = A_2$ are consistent, respectively, and

$$\min_{B_2XC_2=A_2} r(\,A_1 - B_1XC_1\,) = \min_{B_1XC_1=A_1} r(\,A_2 - B_2XC_2\,) = 0.$$

We immediately find from (20.4) the following well-known results.

Corollary 20.3[103][143]. *The pair of matrix equations $B_1XC_1 = A_1$ and $B_2XC_2 = A_2$ have a common solution if and only if $B_1XC_1 = A_1$ and $B_2XC_2 = A_2$ are consistent, respectively, and*

$$r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} = r\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} + r[\,C_1,\ C_2\,].$$

Corollary 20.4. *Suppose that the pair of matrix equations $B_1XC_1 = A_1$ and $B_2XC_2 = A_2$ are consistent, respectively, and denote their solution sets by*

$$\Omega_1 = \{\, X \mid B_1XC_1 = A_1 \,\} \quad \textit{and} \quad \Omega_2 = \{\, X \mid B_2XC_2 = A_2 \,\}.$$

Then

(a) $\Omega_2 \subseteq \Omega_1$ *holds if and only if* $B_1 = 0$ *or* $C_1 = 0$ *or* $r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} = r(B_2) + r(C_2)$.

(b) *Under $B_i \neq 0$ and $C_i \neq 0$, $i = 1, 2$, the two equations $B_1XC_1 = A_1$ and $B_2XC_2 = A_2$ have the same solution set, i.e., $\Omega_1 = \Omega_2$, if and only if*

$$r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} = r(B_1) + r(C_1) = r(B_2) + r(C_2).$$

Another result related to the pair of matrix equations $B_1XC_1 = A_1$ and $B_2XC_2 = A_2$ is given below, which was presented by the author in [139].

Corollary 20.5. *Suppose that $B_1X_1C_1 = A_1$ and $B_2X_2C_2 = A_2$ are consistent, respectively, where X_1 and X_2 have the same size. Then*

$$\min_{\substack{B_1X_1C_1 = A_1 \\ B_2X_2C_2 = A_2}} r(\, X_1 - X_2 \,) = r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} - r\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} - r[\, C_1, \ C_2 \,]. \tag{20.5}$$

Finally we present two results on rank invariance and range invariance of $A_1 - B_1XC_1$ subject to $B_2XC_2 = A_2$.

Theorem 20.6. *Suppose that $B_2XC_2 = A_2$ is consistent. Then the rank of $A_1 - B_1XC_1$ is invariant subject to $B_2XC_2 = A_2$ if and only if*

$$r\begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \\ 0 & B_2 \end{bmatrix} = r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix} + r(B_2) \quad \textit{and} \quad r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} = r\begin{bmatrix} A_1 & B_1 & 0 \\ C_1 & 0 & C_2 \end{bmatrix} + r(B_2),$$

or

$$r\begin{bmatrix} A_1 & B_1 & 0 \\ C_1 & 0 & C_2 \end{bmatrix} = r[\, A_1, \ B_1 \,] + r(C_2) \quad \textit{and} \quad r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} = r\begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \\ 0 & B_2 \end{bmatrix} + r(C_2),$$

or

$$r\begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \\ 0 & B_2 \end{bmatrix} = r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix} + r(B_2) \quad \textit{and} \quad r\begin{bmatrix} A_1 & B_1 & 0 \\ C_1 & 0 & C_2 \end{bmatrix} = r[\, A_1, \ B_1 \,] + r(C_2).$$

Proof. It is obvious that the rank of $A_1 - B_1XC_1$ is invariant subject to $B_2XC_2 = A_2$ if and only if

$$\max_{B_2XC_2 = A_2} r(\, A_1 - B_1XC_1 \,) = \min_{B_2XC_2 = A_2} r(\, A_1 - B_1XC_1 \,).$$

Applying Theorem 20.1 to it produces the desired result in the theorem. □

Theorem 20.7. *Suppose that $B_2XC_2 = A_2$ is consistent with $B_1F_{B_2} \neq 0$ and $C_1E_{C_2} \neq 0$. Then*

(a) *The range $R(A_1 - B_1XC_1)$ is invariant subject to $B_2XC_2 = A_2$ if and only if*

$$r\begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \\ 0 & B_2 \end{bmatrix} = r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix} + r(B_2) \quad \textit{and} \quad r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} = r\begin{bmatrix} A_1 & B_1 & 0 \\ C_1 & 0 & C_2 \end{bmatrix} + r(B_2).$$

(b) *The range $R[(A_1 - B_1XC_1)^T]$ is invariant subject to $B_2XC_2 = A_2$ if and only if*

$$r\begin{bmatrix} A_1 & B_1 & 0 \\ C_1 & 0 & C_2 \end{bmatrix} = r[\, A_1, \ B_1 \,] + r(C_2) \ \textit{and} \ r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} = r\begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \\ 0 & B_2 \end{bmatrix} + r(C_2).$$

Proof. Follows from Theorem 19.4. □

Chapter 21

Extreme ranks of the Schur complement $D - CA^-B$

With the proper background of rank formulas presented in the previous chapter, we are now able to systematically deal with ranks of generalized Schur complements and various related topics. As is well known, for a given block matrix $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ over an arbitrary field $\mathcal{F}$, where A, B, C and D are $m \times n$, $m \times k$, $l \times n$ and $l \times k$ matrices, respectively, a generalized Schur complement of A in M is defined to be

$$S_A = D - CA^-B, \tag{21.1}$$

where A^- is an inner inverse of A, i.e., $A^- \in \{ X \mid AXA = A \}$. As one of the most important matrix expressions in matrix theory, there have been many results in the literature on generalized Schur complements and their applications (see, e.g., [3, 18, 19, 23, 24, 44, 83, 106, 109, 128]). Some of the work focused on equalities and inequalities for ranks of generalized Schur complements. The two rank well-known inequalities (see [19, 23, 83]) related to the Schur complement S_A are given by

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} \geq r(A) + r(\, D - CA^-B\,), \tag{21.2}$$

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} \leq r\begin{bmatrix} A & B \\ C & CA^-B \end{bmatrix} + r(\, D - CA^-B\,). \tag{21.3}$$

Both of them in fact give upper and lower bounds for the rank of the the Schur complement $D - CA^-B$, but they are not, in general, the maximal and the minimal ranks of $D - CA^-B$ with respect to A^-. Note that A^- is in fact a solution of the matrix equation $AXA = A$. Thus Schur complement $D - CA^-B$ may be regarded as a matrix expression $D - CXB$, where X is a solution of the matrix equation $AXA = A$. In that case, applying the rank formulas in Chapter 20, we can simply establish the following.

Theorem 21.2. *Let $S_A = D - CA^-B$ be given by* (21.1). *Then*

(a) *The maximal rank of S_A with respect to A^- is*

$$\max_{A^-} r(\, D - CA^-B\,) = \min\left\{ r[\, C, \ D\,], \quad r\begin{bmatrix} B \\ D \end{bmatrix}, \quad r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r(A) \right\}. \tag{21.4}$$

(b) *The minimal rank of S_A with respect to A^- is*

$$\min_{A^-} r(\, D - CA^-B\,) = r(A) + r[\, C, \ D\,] + r\begin{bmatrix} B \\ D \end{bmatrix} + r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix}. \tag{21.5}$$

Proof. It is quite obvious that

$$\max_{A^-} r(\, D - CA^-B\,) = \max_{AXA=A} r(\, D - CXB\,), \qquad \min_{A^-} r(\, D - CA^-B\,) = \min_{AXA=A} r(\, D - CXB\,).$$

Thus we obtain (21.4) and (21.5) by Theorem 20.1. $\square$

Eq. (21.5) can also be written as

$$\min_{A^-} r(D - CA^- B) = \left(r[C,\ D] + r\begin{bmatrix} B \\ D \end{bmatrix} - r\begin{bmatrix} 0 & B \\ C & D \end{bmatrix} \right)$$
$$+ \left(r\begin{bmatrix} A & B \\ C & D \end{bmatrix} + r\begin{bmatrix} 0 & B \\ C & D \end{bmatrix} + r(A) - r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} \right),$$

and the two quantities in the parentheses on the right hand-side of the above equality are nonnegative.

The two formulas in (21.4) and (21.5) can further simplify when A, B, C and D satisfy some conditions, such as, $R(D) \subseteq R(C)$ and $R(D^T) \subseteq R(B^T)$; $R(D) \cap R(C) = \{0\}$ and $R(D^T) \cap R(B^T) = \{0\}$; $R(C) \subseteq R(D)$ and $R(B^T) \subseteq R(D^T)$. The reader can easily list the corresponding results.

Corollary 21.3. *The rank of $D - CA^- B$ is invariant with respect to the choice of A^- if and only if*

$$r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} = r\begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad \text{and} \quad r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} = r\begin{bmatrix} B \\ D \end{bmatrix} + r(A),$$

or

$$r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} = r[C,\ D] + r(A) \quad \text{and} \quad r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} = r\begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

or

$$r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} = r[C,\ D] + r(A) \quad \text{and} \quad r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} = r\begin{bmatrix} B \\ D \end{bmatrix} + r(A).$$

Proof. It is obvious that the rank of $D - CA^- B$ is invariant with respect to the choice of A^- if and only if

$$\max_{A^-} r(D - CA^- B) = \min_{A^-} r(D - CA^- B).$$

Applying Theorem 20.6 to it leads to the desired result in the corollary. $\square$

Corollary 21.4. *Let S_A be given by* (21.1) *with $E_A B \neq 0$ and $CF_A \neq 0$.*

(a) *The range $R(D - CA^- B)$ is invariant with respect to the choice of A^- if and only if*

$$r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} = r\begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad \text{and} \quad r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} = r\begin{bmatrix} B \\ D \end{bmatrix} + r(A).$$

(b) *The range $R[(D - CA^- B)^T]$ is invariant with respect to the choice of A^- if and only if*

$$r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} = [C,\ D] + r(A) \quad \text{and} \quad r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} = r\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Proof. Follows from Theorem 20.7. $\square$

Combining (21.2) and (21.4), (21.3) and (21.5), we derive the following several results.

Theorem 21.5. *Let S_A be given by* (21.1). *Then*

(a) *There is an $A^- \in \{A^-\}$ such that*

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r(A) + r(D - CA^- B), \tag{21.6}$$

if and only if

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} \leq r(A) + \min\left\{ r[C,\ D], \quad r\begin{bmatrix} B \\ D \end{bmatrix} \right\}. \tag{21.7}$$

(b) *The equality* (21.6) *holds for all* $A^- \in \{A^-\}$ *if and only if*

$$r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} = r(A) + r[\, C, \ D\,] \quad and \quad r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} = r(A) + r\begin{bmatrix} B \\ D \end{bmatrix}. \tag{21.8}$$

Proof. Note from (21.2) that $r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r(A)$ is an upper bound for $r(\, D - CA^-B\,)$. Thus there is an $A^- \in \{A^-\}$ such that (21.6) holds if and only if

$$\max_{A^-} r(\, D - CA^-B\,) = r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r(A).$$

Putting (21.4) in it immediately yields (21.7). On the other hand, (21.6) holds for all $A^- \in \{A^-\}$ if and only if

$$\min_{A^-} r(\, D - CA^-B\,) = r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r(A).$$

Putting (21.5) in it yields (21.8). $\square$

The rank equality (21.6) was examined by Carlson in [23] and Marsaglia and Styan in [83]. Their conclusion is that (21.6) holds if and only if $(\, I - AA^-\,)B(\, I - S_A^-S_A\,) = 0$, $(\, I - S_A^-S_A\,)C(\, I - A^-A\,) = 0$ and $(\, I - AA^-\,)BS_A^-C(\, I - A^-A\,) = 0$. In comparison, (21.7) has no inner inverses in it, thus it is simpler and is easier to verify.

Theorem 21.6. *Let* S_A *be given by* (21.1). *Then*

(a) *There is an* $A^- \in \{A^-\}$ *such that*

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r\begin{bmatrix} A & B \\ C & CA^-B \end{bmatrix} + r(\, D - CA^-B\,), \tag{21.9}$$

if and only if

$$r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} = r[\, A, \ B\,] + r[\, C, \ D\,] \quad and \quad r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r\begin{bmatrix} B \\ D \end{bmatrix}. \tag{21.10}$$

(b) *The equality* (21.9) *holds for all* $A^- \in \{A^-\}$ *if and only if*

$$R(B) \subseteq R(A) \quad and \quad R(C^T) \subseteq R(A^T), \tag{21.11}$$

or

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r\begin{bmatrix} B \\ D \end{bmatrix} \quad and \quad R(B) \subseteq R(A), \tag{21.12}$$

or

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r[\, A, \ B\,] + r[\, C, \ D\,] \quad and \quad R(C^T) \subseteq R(A^T). \tag{21.13}$$

Proof. Note from (21.3) that $r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A & B \\ C & CA^-B \end{bmatrix}$ is a lower bound for $r(\, D - CA^-B\,)$. Thus (21.9) holds if and only if

$$\min_{A^-} r(\, D - CA^-B\,) = r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A & B \\ C & CA^-B \end{bmatrix} = r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A \\ C \end{bmatrix} - r[\, A, \ B\,] + r(A).$$

Combining it with (21.5) yields

$$r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} + r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r\begin{bmatrix} B \\ D \end{bmatrix} + r[\, A, \ B\,] + r[\, C, \ D\,],$$

which is obviously equivalent to (21.10). On the other hand, (21.9) holds for all $A^- \in \{A^-\}$ if and only if

$$\max_{A^-} r(\, D - CA^-B\,) = r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A & B \\ C & CA^-B \end{bmatrix}.$$

Combining it with (21.4) yields (21.11)—(21.13). □

As a special case of Schur complements, the rank and the range of the product CA^-B and their applications were examined by Baksalary and Kala in [6], Baksalary and Mathew in[7] and Gross in [53]. Based on the previous several theorems and corollaries, we now have the following three corollaries.

Corollary 21.7. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ and $C \in \mathcal{F}^{l\times n}$ be given. Then*

(a) *The maximal rank of CA^-B with respect to A^- is*

$$\max_{A^-} r(CA^-B) = \min\left\{ r(B), \quad r(C), \quad r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r(A) \right\}. \tag{21.14}$$

(b) *The minimal rank of CA^-B with respect to A^- is*

$$\min_{A^-} r(CA^-B) = r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r\begin{bmatrix} A \\ C \end{bmatrix} - r[\, A, \ B\,] + r(A). \tag{21.15}$$

(c) *There is an $A^- \in \{A^-\}$ such that $CA^-B = 0$ if and only if*

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A, \ B\,] - r(A). \tag{21.16}$$

(d) *$CA^-B = 0$ holds for all $A^- \in \{A^-\}$ if and only if $B = 0$ or $C = 0$ or $r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(A)$.*

(e)[53] *The rank of CA^-B is invariant with respect to the choice of A^- if and only if*

$$R(B) \subseteq R(A) \quad \text{and} \quad R(C^T) \subseteq R(A^T), \tag{21.17}$$

or

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r[\, A, \ B\,] + r(C) \quad \text{and} \quad R(C^T) \subseteq R(A^T), \tag{21.18}$$

or

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r(B) \quad \text{and} \quad R(B) \subseteq R(A). \tag{21.19}$$

Corollary 21.8. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ and $C \in \mathcal{F}^{l\times n}$ be given with $B \neq 0$ and $C \neq 0$. Then*

(a)[53] *The range $R(CA^-B)$ is invariant with respect to the choice of A^- if and only if $R(B) \subseteq R(A)$ and $R(C^T) \subseteq R(A^T)$, or*

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r[\, A, \ B\,] + r(C) \quad \text{and} \quad R(C^T) \subseteq R(A^T). \tag{21.20}$$

(b)[53] *The range $R[(CA^-B)^T]$ is invariant with respect to the choice of A^- if and only if $R(B) \subseteq R(A)$ and $R(C^T) \subseteq R(A^T)$, or*

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r(B) \quad \text{and} \quad R(B) \subseteq R(A). \tag{21.21}$$

(c)[7] *The rank of CA^-B is invariant with respect to the choice of A^- if and only if $R(CA^-B)$ or $R[(CA^-B)^T]$ is invariant with respect to the choice of A^-.*

Corollary 21.9. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ and $C \in \mathcal{F}^{l\times n}$. Then*

$$\max_{A^-} r(A^-B) = r(B), \qquad \max_{B^-} r(B^-A) = r(A), \tag{21.22}$$

$$\max_{A^-} r(AA^-B) = \max_{B^-} r(BB^-A) = \min\{\, r(A), \quad r(B)\,\}, \tag{21.23}$$

$$\max_{A^-} r(CA^-) = r(C), \qquad \max_{C^-} r(AC^-) = r(A), \tag{21.24}$$

$$\max_{A^-} r(CA^-A) = \max_{C^-} r(AC^-C) = \min\{\, r(A), \quad r(C)\,\}, \tag{21.25}$$

$$\min_{A^-} r(AA^-B) = \min_{B^-} r(BB^-A) = \min_{A^-} r(A^-B) = \min_{B^-} r(B^-A) = r(A) + r(B) - r[\, A, \ B\,], \tag{21.26}$$

$$\min_{A^-} r(CA^-A) = \min_{C^-} r(AC^-C) = \min_{A^-} r(CA^-) = \min_{C^-} r(AC^-) = r(A) + r(C) - r\begin{bmatrix} A \\ C \end{bmatrix}. \tag{21.27}$$

In particular,

(a) *There are* A^- *and* B^- *such that* $A^-B = 0$ *and* $B^-A = 0$ *if and only if* $R(A) \cap R(B) = \{0\}$.

(b) *There are* A^- *and* C^- *such that* $CA^- = 0$ *and* $AC^- = 0$ *if and only if* $R(A^T) \cap R(C^T) = \{0\}$.

The two formulas in (21.4) and (22.5) can help to establish various rank equalities for matrix expressions that involve inner inverses of matrices, and then to derive from them various consequences. Below are some of them.

Theorem 21.10. *Let* $A, B \in \mathcal{F}^{m\times n}$ *be given. Then*

$$\max_{B^-} r(\, A - AB^-A\,) = \min\{\ r(A), \quad r(\, B - A\,) - r(B) + r(A)\ \}, \tag{21.28}$$

$$\min_{B^-} r(\, A - AB^-A\,) = \min_{A^-,\, B^-} r(\, A^- - B^-\,) = r(\, A - B\,) + r(A) + r(B) - r[\, A, \ B\,] - r\begin{bmatrix} A \\ B \end{bmatrix}. \tag{21.29}$$

In particular,

(a) *A and B have a common inner inverse if and only if* $r(\, A - B\,) = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A, \ B\,] - r(A) - r(B)$.

(b) *The inclusion* $\{B^-\} \subseteq \{A^-\}$ *holds if and only if* $A = 0$ *or* $r(\, B - A\,) = r(B) - r(A)$.

(c)[103] $\{A^-\} = \{B^-\}$ *holds if and only if* $A = B$.

(d) $\{A^-\} \cap \{B^-\} = \emptyset$ *holds if and only if* $r(\, A - B\,) > r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A, \ B\,] - r(A) - r(B)$.

(e) *If* $R(A) \cap R(B) = \{0\}$ *and* $R(A^T) \cap R(B^T) = \{0\}$, *then there must exist* $A^- \in \{A^-\}$ *and* $B^- \in \{B^-\}$ *such that* $A^- = B^-$.

Proof. Eq. (21.28) follows from (21.4); (21.29) follows from (21.5) and (20.5). The results in Parts (a)—(e) are direct consequences of (21.28) and (21.29). □

A lot of consequences can be derived from Theorem 21.10. For example, let $B = A^k$ in (21.29). Then we get

$$\min_{A^-,\, (A^k)^-} r[\, A^- - (A^k)^-\,] = r(\, A - A^k\,) + r(A^k) - r(A). \tag{21.30}$$

Thus A and A^k have a common inner inverse if and only if $r(\, A - A^k\,) = r(A) - r(A^k)$. In that case, $\{A^-\} \subseteq \{(A^k)^-\}$ holds by Theorem 21.10(b).

Replacing A and B in (21.29) by $I_m - A$ and A, respectively, we can get by (1.16)

$$\min_{(I_m - A)^-,\, A^-} r[\, (I_m - A)^- - A^-\,] = r(I_m - 2A) + r(I_m - A) + r(A) - 2m = r[\, A(I_m - A)(I_m - 2A)\,]. \tag{21.31}$$

Thus $I_m - A$ and A have a common inner inverse if and only if $A(I_m - A)(I_m - 2A) = 0$.

Replacing A and B in (21.29) by $A - I_m$ and A, respectively, we can get

$$\min_{(A - I_m)^-,\, A^-} r[\, (A - I_m)^- - A^-\,] = r(I_m - A) + r(A) - m = r(\, A^2 - A\,). \tag{21.32}$$

Thus A is idempotent if and only if $A - I_m$ and A have a common inner inverse, this fact could be regarded as a new characterization of idempotent matrix.

Replacing A and B in (21.29) by $I_m + A$ and A, respectively, we can get by (1.11)

$$\min_{(I_m+A)^{-},\, A^{-}} r[\,(I_m + A)^{-} - A^{-}\,] = r(I_m + A) + r(A) - m = r(A + A^2). \tag{21.33}$$

Thus $I_m + A$ and A have a common inner inverse if and only if $A^2 = -A$.

Replacing A and B in (21.29) by $I_m + A$ and $I_m - A$, respectively, we can get by (1.15)

$$\min_{(I_m+A)^{-},\, (I_m-A)^{-}} r[\,(I_m + A)^{-} - (I_m - A)^{-}\,] = r(A) + r(I_m + A) + r(I_m - A) - 2m = r(\,A^3 - A\,). \tag{21.34}$$

In particular, $I_m + A$ and $I_m - A$ have a common inner inverse if and only if A is tripotent.

Replacing A and B in (21.15) by $A + I_m$ and $A - I_m$, respectively, we can get by (1.12)

$$\min_{(A+I_m)^{-},\, (A-I_m)^{-}} r[\,(A + I_m)^{-} - (A - I_m)^{-}\,] = r(A + I_m) + r(A - I_m) - m = r(\,A^2 - I_m\,). \tag{21.35}$$

This implies that A is involutory if and only if $A + I_m$ and $A - I_m$ have a common inner inverse, this fact could be regarded as a new characterization of involutory matrix.

Now suppose $\lambda_1 \neq \lambda_2$ are two scalars. Then it is easy to show by (21.29) and (1.16) the following two rank equalities

$$\min_{(\lambda_1 I_m-A)^{-},\, (\lambda_2 I_m-A)^{-}} r[\,(\lambda_1 I_m - A)^{-} - (\lambda_2 I_m - A)^{-}\,] = r[\,(\lambda_1 I_m - A)(\lambda_2 I_m - A)\,], \tag{21.36}$$

$$\min_{(I_m-\lambda_1 A)^{-},\, (I_m-\lambda_2 A)^{-}} r[\,(I_m - \lambda_1 A)^{-} - (I_m - \lambda_2 A)^{-}\,] = r[\,A(I_m - \lambda_1 A)(I_m - \lambda_2 A)\,]. \tag{21.37}$$

Thus the two matrices $\lambda_1 I_m - A$ and $\lambda_2 I_m - A$ have a common inner inverse if and only if $(\lambda_1 I_m - A)(\lambda_2 I_m - A) = 0$. The two matrices $I_m - \lambda_1 A$ and $I_m - \lambda_2 A$ have have a common inner inverse if and only if $A(I_m - \lambda_1 A)(I_m - \lambda_2 A) = 0$.

Again replacing A and B in (21.15) by $A^k + A$ and $A^k - A$, respectively, we can get by (1.14)

$$\min_{(A^k+A)^{-},\, (A^k-A)^{-}} r[\,(A^k + A)^{-} - (A^k - A)^{-}\,] = r(A^k + A) + r(A^k - A) - r(A) = r(\,A^{2k-1} - A\,). \tag{21.38}$$

In particular, $A^{2k-1} = A$ if and only if $A^k + A$ and $A^k - A$ have a common inner inverse.

In general, suppose that $p(x)$ and $q(x)$ are two polynomials without common roots. Then there is

$$\min_{p^{-}(A),\, q^{-}(A)} r[\,p^{-}(A) - q^{-}(A)\,] = r[\,p^2(A)q(A) - p(A)q^2(A)\,]. \tag{21.39}$$

Thus $p(A)$ and $q(A)$ have a common inner inverse if and only if $p^2(A)q(A) = p(A)q^2(A)$.

¿From (21.29) we also get

$$\min_{(A+B)^{-},\, A^{-}} r[\,(\,A + B\,)^{-} - A^{-}\,] \tag{21.40}$$

$$= \min_{(A+B)^{-},\, B^{-}} r[\,(\,A + B\,)^{-} - B^{-}\,] = r(\,A + B\,) + r(A) + r(B) - r[\,A,\ B\,] - r\begin{bmatrix} A \\ B \end{bmatrix}. \tag{21.41}$$

Hence we see that if $R(A) \cap R(B) = \{0\}$ and $R(A^T) \cap R(B^T) = \{0\}$, then $A + B$ and A must have a common inner inverse, meanwhile then $A + B$ and B must have a common inner inverse.

Now let $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$ and $N = \begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix}$. Then we get from (21.29) that

$$\min_{M^{-},\, N^{-}} r(\,M^{-} - N^{-}\,) = r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r\begin{bmatrix} A \\ C \end{bmatrix} - r[\,A,\ B\,] + r(A). \tag{21.42}$$

Hence M and N have a common inner inverse if and only if

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,] - r(A). \tag{21.43}$$

Next let $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$ and $N = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$. Then we can also get from (21.29) that

$$\min_{M^-,\, N^-} r(\, M^- - N^- \,) = r \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r \begin{bmatrix} A \\ C \end{bmatrix} - r[\, A, \ B\,] + r(A). \tag{21.44}$$

Hence M and N have a common inner inverse also if and only if (21.43) holds.

Furthermore let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ and $N = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$. Then we can also derive from (21.29) that

$$\begin{aligned} &\min_{M^-,\, N^-} r(\, M^- - N^- \,) \\ &= r(A) + r(B) + r(C) + r(D) + r \begin{bmatrix} A & B \\ C & D \end{bmatrix} - r[\, A, \ B\,] - r[\, C, \ D\,] - r \begin{bmatrix} A \\ C \end{bmatrix} - r \begin{bmatrix} B \\ D \end{bmatrix}. \end{aligned} \tag{21.45}$$

Hence M and N have a common inner inverse if and only if

$$r \begin{bmatrix} A & B \\ C & D \end{bmatrix} = r \begin{bmatrix} A \\ C \end{bmatrix} + r \begin{bmatrix} B \\ D \end{bmatrix} + r[\, A, \ B\,] + r[\, C, \ D\,] - r(A) - r(B) - r(C) - r(D) \tag{21.46}$$

holds. In particular, if $r(M) = r(A) + r(B) + r(C) + r(D)$, then M and N must have a common inner inverse.

Finally let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ and $N = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$. Then we derive from (21.29) that

$$\min_{M^-,\, N^-} r(\, M^- - N^- \,) = r \begin{bmatrix} A & B \\ C & D \end{bmatrix} + r \begin{bmatrix} 0 & B \\ C & D \end{bmatrix} + r(A) - r \begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} - r \begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix}. \tag{21.47}$$

Hence M and N have a common inner inverse if and only if

$$r \begin{bmatrix} A & B \\ C & D \end{bmatrix} = r \begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} + r \begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} - r \begin{bmatrix} 0 & B \\ C & D \end{bmatrix} - r(A). \tag{21.48}$$

Theorem 21.11. *Let $A, B \in \mathcal{F}^{m\times n}$ be given. Then*

$$\begin{aligned} &\max_{(A+B)^-} r[\, A(\, A+B\,)^- B\,] \\ &= \max_{(A+B)^-} r[\, B(\, A+B\,)^- A\,] = \min\{\, r(A), \ r(B), \ r(A) + r(B) - r(A+B)\,\}, \end{aligned} \tag{21.49}$$

and

$$\begin{aligned} &\min_{(A+B)^-} r[\, A(\, A+B\,)^- B\,] \\ &= \min_{(A+B)^-} r[\, B(\, A+B\,)^- A\,] = r(\, A+B\,) + r(A) + r(B) - r[\, A, \ B\,] - r \begin{bmatrix} A \\ B \end{bmatrix}. \end{aligned} \tag{21.50}$$

In particular,

(a) *There is an $(\, A+B\,)^-$ such that $A(\, A+B\,)^- B = 0$ if and only if $r(\, A+B\,) = r \begin{bmatrix} A \\ B \end{bmatrix} + r[\, A, \ B\,] - r(A) - r(B)$.*

(b) *$A(\, A+B\,)^- B = 0$ holds for all $(\, A+B\,)^-$ if and only if $A = 0$ or $B = 0$ or $r(\, A+B\,) = r(A) + r(B)$.*

(c)[117] *The rank of $A(\, A+B\,)^- B$ is invariant with respect to the choice of $(\, A+B\,)^-$ if and only if $R(B) \subseteq R(\, A+B\,)$ and $R(A^T) \subseteq R(\, A^T + B^T\,)$, that is, A and B are parallel summable.*

A parallel result to Theorem 21.11 is

Theorem 21.12. *Let $A, B \in \mathcal{F}^{m\times n}$ be given. Then*

$$\max_{A^-,\, B^-} r[\, A^-(\, A+B\,)B^-\,] = \max_{A^-,\, B^-} r[\, B^-(\, A+B\,)A^-\,] = r(\, A+B\,), \tag{21.51}$$

and

$$\begin{aligned} &\min_{A^-,\, B^-} r[\, A^-(\, A + B\,)B^-\,] \\ &= \min_{A^-,\, B^-} r[\, B^-(\, A + B\,)A^-\,] = r(\, A + B\,) + r(A) + r(B) - r[\, A,\ B\,] - r\begin{bmatrix} A \\ B \end{bmatrix}. \end{aligned} \tag{21.52}$$

In particular,

(a) *There are A^-, B^- such that $A^-(\, A + B\,)A^- = 0$ if and only if $r(\, A + B\,) = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A,\ B\,] - r(A) - r(B)$.*

(b) *The rank of $A^-(\, A + B\,)B^-$ is invariant with respect to the choice of A^-, B^- if and only if $R(A) = R(B)$ and $R(A^T) = R(B^T)$.*

Proof. According to (21.22) and (21.26) we first get

$$\max_{A^-} r[\, A^-(\, A + B\,)B^-\,] = r[\,(\, A + B\,)B^-\,],$$

$$\begin{aligned} \min_{A^-} r[\, A^-(\, A + B\,)B^-\,] &= r(A) + r[\,(\, A + B\,)B^-\,] - r[\, A,\ (\, A + B\,)B^-\,] \\ &= r(A) + r[\,(\, A + B\,)B^-\,] - r[\, A,\ B\,] \end{aligned}$$

Next by (21.24) and (21.27), we find

$$\max_{B^-} r[\,(\, A + B\,)B^-\,] = r(\, A + B\,),$$

$$\min_{B^-} r[\,(\, A + B\,)B^-\,] = r(B) + r(\, A + B\,) - r\begin{bmatrix} A + B \\ B \end{bmatrix} = r(B) + r(\, A + B\,) - r\begin{bmatrix} A \\ B \end{bmatrix}.$$

Combining them yields (21.51) and (21.52). □.

From (21.52) we can also find some interesting consequences. For example, let $B = I_m - A$ in (21.52), we can get

$$\min_{A^-,\, (I_m - A)^-} r[\, A^-(I_m - A)^-\,] = \min_{A^-,\, (I_m - A)^-} r[\, A^-(I_m - A)^-\,] = r(\, A - A^2\,). \tag{21.53}$$

Thus A is idempotent if and only if there are A^- and $(I_m - A)^-$ such that $A^-(I_m - A)^- = 0$, which could be regard as a new characterization for idempotent matrix.

Replacing A and B in (21.52) by $I_m + A$ and $I_m - A$, respectively, we can get

$$\min_{(I_m + A)^-,\, (I_m - A)^-} r[\,(I_m + A)^-(I_m - A)^-\,] = \min_{(I_m + A)^-,\, (I_m - A)^-} r[\,(I_m - A)^-(I_m + A)^-\,] = r(\, I_m - A^2\,). \tag{21.54}$$

Thus A is involutory if and only if there are $(I_m + A)^-$ and $(I_m - A)^-$ such that $(I_m + A)^-(I_m - A)^- = 0$, which could be regard as a new characterization for involutory matrix.

In general replacing A and B in (21.52) by $\lambda_1 I_m - A$ and $-(\lambda_2 I_m - A)$, respectively, where $\lambda_1 \neq \lambda_2$, we can get

$$\min_{(\lambda_1 I_m - A)^-,\, (\lambda_2 I_m - A)^-} r[\,(\lambda_1 I_m - A)^-(\lambda_2 I_m - A)^-\,] = r[\,(\lambda_1 I_m - A)(\lambda_2 I_m - A)\,]. \tag{21.55}$$

Thus there are $(\lambda_1 I_m - A)^-$ and $(\lambda_2 I_m - A)^- = 0$ such that $(\lambda_1 I_m - A)^-(\lambda_2 I_m - A)^- = 0$ if and only if $(\lambda_1 I_m - A)(\lambda_2 I_m - A) = 0$.

Motivated by (21.55), we find the the following general result.

Theorem 21.13. *Let $A \in \mathcal{F}^{m \times n}$ and $\lambda_1, \cdots, \lambda_k \in \mathcal{F}$ with $\lambda_i \neq \lambda_j$ for all $i \neq j$. Then*

$$\min_{(\lambda_1 I_m - A)^-,\, \cdots,\, (\lambda_k I_m - A)^-} r[\,(\lambda_1 I_m - A)^- \cdots (\lambda_k I_m - A)^-\,] = r[\,(\lambda_1 I_m - A) \cdots (\lambda_k I_m - A)\,]. \tag{21.56}$$

Proof. According to (21.26) we first get

$$\min_{(\lambda_1 I_m - A)^-} r[\,(\lambda_1 I_m - A)^- \cdots (\lambda_k I_m - A)^-\,]$$
$$= r(\lambda_1 I_m - A) + r[\,(\lambda_2 I_m - A)^- \cdots (\lambda_k I_m - A)^-\,] - r[\,(\lambda_1 I_m - A),\ (\lambda_2 I_m - A)^- \cdots (\lambda_k I_m - A)^-\,]. \tag{21.57}$$

Notice that $\lambda_i \neq \lambda_j$ for $i \neq j$. Then there must be

$$r[\,(\lambda_1 I_m - A),\ (\lambda_2 I_m - A)^- \cdots (\lambda_k I_m - A)^-\,] = m \tag{21.58}$$

for all $(\lambda_2 I_m - A)^-, \cdots,\ (\lambda_k I_m - A)^-$. To show this fact, we need the following two rank formulas

$$\min_{B^-} r[\,A,\ B^-\,] = r(A) + r(B) - r(BA), \tag{21.59}$$
$$\min_{B^-} r[\,A,\ B^-C\,] = r(A) + r(B) - r(BA) + r[\,BA,\ C\,] - r[\,B,\ C\,], \tag{21.60}$$

We see first by (21.59) and (1.16) that for $\lambda_t \neq \lambda_j,\ j = 1, \cdots, i$ there is

$$\min_{(\lambda_t I_m - A)^-} r[\,(\lambda_1 I_m - A) \cdots (\lambda_i I_m - A),\ (\lambda_t I_m - A)^-\,]$$
$$= r[\,(\lambda_1 I_m - A) \cdots (\lambda_i I_m - A)\,] + r(\lambda_t I_m - A) - r[\,(\lambda_1 I_m - A) \cdots (\lambda_i I_m - A)(\lambda_t I_m - A)\,] = m.$$

That is to say,

$$r[\,(\lambda_1 I_m - A) \cdots (\lambda_i I_m - A),\ (\lambda_t I_m - A)^-\,] = m$$

holds for any $(\lambda_t I_m - A)^-$ with $\lambda_t \neq \lambda_j,\ j = 1, \cdots, i$. Now suppose that

$$r[\,(\lambda_1 I_m - A) \cdots (\lambda_i I_m - A),\ (\lambda_{t+1} I_m - A)^- \cdots (\lambda_k I_m - A)^-\,] = m$$

holds for all $(\lambda_{t+1} I_m - A)^-, \cdots, (\lambda_k I_m - A)^-$ and $1 \leq i < t < k$. Then we can obtain by (21.60), (1.16) and induction hypothesis that

$$\min_{(\lambda_t I_m - A)^-} r[\,(\lambda_1 I_m - A) \cdots (\lambda_i I_m - A),\ (\lambda_t I_m - A)^- \cdots (\lambda_k I_m - A)^-\,]$$
$$= r[\,(\lambda_1 I_m - A) \cdots (\lambda_i I_m - A)\,] + r(\lambda_t I_m - A) - r[\,(\lambda_1 I_m - A) \cdots (\lambda_i I_m - A)(\lambda_t I_m - A)\,]$$
$$\quad + r[\,(\lambda_1 I_m - A) \cdots (\lambda_i I_m - A)(\lambda_t I_m - A),\ (\lambda_{t+1} I_m - A)^- \cdots (\lambda_k I_m - A)^-\,]$$
$$\quad - r[\,(\lambda_t I_m - A),\ (\lambda_{t+1} I_m - A)^- \cdots (\lambda_k I_m - A)^-\,]$$
$$= m + m - m = m,$$

that is,

$$r[\,(\lambda_1 I_m - A) \cdots (\lambda_i I_m - A),\ (\lambda_t I_m - A)^- \cdots (\lambda_k I_m - A)^-\,] = m \tag{21.61}$$

holds for all $(\lambda_t I_m - A)^-, \cdots, (\lambda_k I_m - A)^-$ and $1 \leq i < t < k$. When $i = 1$ and $t = 2$, (21.61) becomes (21.58). In that case, (21.57) reduces to

$$\min_{(\lambda_1 I_m - A)^-} r[\,(\lambda_1 I_m - A)^- \cdots (\lambda_k I_m - A)^-\,] = r(\lambda_1 I_m - A) + r[\,(\lambda_2 I_m - A)^- \cdots (\lambda_k I_m - A)^-\,] - m. \tag{21.62}$$

Repeatedly applying (21.62) for the product $(\lambda_2 I_m - A)^- \cdots (\lambda_k I_m - A)^-$ in (21.62), we eventually get

$$\min_{(\lambda_1 I_m - A)^-, \cdots, (\lambda_k I_m - A)^-} r[\,(\lambda_1 I_m - A)^- \cdots (\lambda_k I_m - A)^-\,] = r(\lambda_1 I_m - A) + \cdots + r(\lambda_k I_m - A) - m(k-1),$$

which, by (1.16), is the desired formula (21.56). □

When two square matrices A and B of the same size are nonsingular, it is well known that $A^{-1} + B^{-1} = A^{-1}(A+B)B^{-1}$. This fact motivates us to consider the relationship between $A^- + B^-$ and $A^-(A+B)B^-$ in general case. Using the rank formula (21.5) we can simply find that

$$\min_{A^-,\, B^-} r[\,A^- + B^- - A^-(A+B)B^-\,] = 0,$$

which implies the following.

Theorem 21.14. *Let $A,\ B \in \mathcal{F}^{m\times n}$ be given. Then there must exist A^- and B^- such that*

$$A^- + B^- = A^-(A+B)B^- \tag{21.63}$$

holds.

As applications, we can simply get from (21.63) that there must exist A^- and $(I_m - A)^-$ such that

$$A^- + (I_m - A)^- = A^-(I_m - A)^-. \tag{21.64}$$

and there must exist $A(I_m + A)^-$ and $(I_m - A)^-$ such that

$$(I_m + A)^- + (I_m - A)^- = 2(I_m + A)^-(I_m - A)^-. \tag{21.65}$$

This result leads to the following conjecture .

Conjecture 21.15. *Let $A \in \mathcal{F}^{m\times n}$ and $\lambda_1, \cdots, \lambda_k \in \mathcal{F}$ with $\lambda_i \neq \lambda_j$. Then there exist $(\lambda_1 I_m - A)^-, \cdots, (\lambda_k I_m - A)^-$ such that*

$$\frac{1}{p_1}(\lambda_1 I_m - A)^- + \cdots + \frac{1}{p_k}(\lambda_k I_m - A)^- = (\lambda_1 I_m - A)^- \cdots (\lambda_k I_m - A)^-. \tag{21.66}$$

where

$$p_i = (\lambda_1 - \lambda_i)\cdots(\lambda_{i-1} - \lambda_i)(\lambda_{i+1} - \lambda_i)\cdots(\lambda_k - \lambda_i), \quad i = 1, \cdots, k.$$

Theorem 21.16. *Let $A,\ B \in \mathcal{F}^{m\times n}$ be given and let $M = \mathrm{diag}(\,A,\ B\,)$ and $N = A + B$. Then*

$$\max_{N^-} r\left(M - \begin{bmatrix} A \\ B \end{bmatrix} N^-[\,A,\ B\,]\right) = \min_{N^-} r\left(M - \begin{bmatrix} A \\ B \end{bmatrix} N^-[\,A,\ B\,]\right) = r(A) + r(B) - r(N). \tag{21.67}$$

That is, the rank of $M - \begin{bmatrix} A \\ B \end{bmatrix} N^-[\,A,\ B\,]$ is invariant with respect to the choice of N^-. In general, for $A_1,\ A_2, \cdots, A_k \in \mathcal{F}^{m\times n}$, there is

$$\begin{aligned} \max_{N^-} r\left(M - \begin{bmatrix} A_1 \\ \vdots \\ A_k \end{bmatrix} N^-[\,A_1,\ \cdots,\ A_k\,]\right) &= \min_{N^-} r\left(M - \begin{bmatrix} A_1 \\ \vdots \\ A_k \end{bmatrix} N^-[\,A_1,\ \cdots,\ A_k\,]\right) \\ &= r(A_1) + \cdots + r(A_k) - r(N), \end{aligned}$$

where $M = \mathrm{diag}(\,A_1, \cdots, A_k\,)$ and $N = A_1 + \cdots + A_k$. In particular, the equality

$$\begin{bmatrix} A_1 \\ \vdots \\ A_k \end{bmatrix} (\,A_1 + \cdots + A_k\,)^-[\,A_1,\ \cdots,\ A_k\,] = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix} \tag{21.68}$$

holds for all $(\,A_1 + \cdots + A_k\,)^-$ if and only if $r(\,A_1 + \cdots + A_k\,) = r(A_1) + \cdots + r(A_k)$.

Theorem 21.17. *Let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ be a partitioned matrix over $\mathcal{F}$. Then*

$$\max_{A^-} r\left(M - \begin{bmatrix} A \\ C \end{bmatrix} A^-[\,A,\ B\,]\right) = \min_{A^-} r\left(M - \begin{bmatrix} A \\ C \end{bmatrix} A^-[\,A,\ B\,]\right) = r(M) - r(A).$$

That is,

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} = r(A) + r\left(M - \begin{bmatrix} A \\ C \end{bmatrix} A^-[\,A,\ B\,]\right),$$

which is exactly the formula (1.5).

Theorem 21.18. *Let* $A \in \mathcal{F}^{m\times m}$ *be given. Then*

$$\max_{A^-} r(\, AA^- - A^-A\,) = \min\{\, 2m - 2r(A), \quad 2r(A)\,\}, \tag{21.69}$$

$$\min_{A^-} r(\, AA^- - A^-A\,) = 2r(A) - 2r(A^2). \tag{21.70}$$

Proof. Since both AA^- and A^-A are idempotent, we see by (3.1) that the rank of $AA^- - A^-A$ can be written as

$$r(\, AA^- - A^-A\,) = r\begin{bmatrix} AA^- \\ A^-A \end{bmatrix} + r[\, AA^-, \ A^-A\,] - r(AA^-) - r(A^-A).$$

Note that $r(AA^-) = r(A^-A) = r(A)$, $r\begin{bmatrix} AA^- \\ A^-A \end{bmatrix} = r\begin{bmatrix} AA^- \\ A \end{bmatrix}$ and $r[\, AA^-, \ A^-A\,] = r[\, A, \ A^-A\,]$. Then

$$r(\, AA^- - A^-A\,) = r\begin{bmatrix} AA^- \\ A \end{bmatrix} + r[\, A, \ A^-A\,] - 2r(A).$$

On the other hand, form the general expression of $A^- = A^\sim + F_AV + WE_A$, we also know that $AA^- = AA^\sim + AWE_A$, and $A^-A = A^\sim A + F_AVA$. Thus AA^- and A^-A are in fact two independent matrix expressions. In that case, we see that

$$\max_{A^-} r(\, AA^- - A^-A\,) = \max_{A^-} r\begin{bmatrix} AA^- \\ A \end{bmatrix} + \max_{A^-} r[\, A, \ A^-A\,] - 2r(A), \tag{21.71}$$

$$\min_{A^-} r(\, AA^- - A^-A\,) = \min_{A^-} r\begin{bmatrix} AA^- \\ A \end{bmatrix} + \min_{A^-} r[\, A, \ A^-A\,] - 2r(A). \tag{21.72}$$

According to (21.4) and (21.5), we easily find that

$$\max_{A^-} r\begin{bmatrix} AA^- \\ A \end{bmatrix} = \max_{A^-} r\left(\begin{bmatrix} 0 \\ A \end{bmatrix} + \begin{bmatrix} A \\ 0 \end{bmatrix} A^- \right) = \min\{\, 2r(A), \quad m\,\},$$

$$\min_{A^-} r\begin{bmatrix} AA^- \\ A \end{bmatrix} = \min_{A^-} r\left(\begin{bmatrix} 0 \\ A \end{bmatrix} + \begin{bmatrix} A \\ 0 \end{bmatrix} A^- \right) = 2r(A) - r(A^2),$$

$$\max_{A^-} r[\, A, \ A^-A\,] = \max_{A^-} r(\, [\, A, \ 0\,] + A^-[\, 0, \ A\,]\,) = \min\{\, 2r(A), \quad m\,\},$$

$$\min_{A^-} r[\, A, \ A^-A\,] = \min_{A^-} r(\, [\, A, \ 0\,] + A^-V[\, 0, \ A\,]\,) = 2r(A) - r(A^2).$$

Putting the above four results in (21.71) and (21.72) yields (21.69) and (21.70). $\square$

Corollary 21.19. *Let* $A \in \mathcal{F}^{m\times m}$ *be given.*

(a) *There is an* A^- *such that* $AA^- - A^-A$ *is nonsingular if and only if* m *is even and* $r(A) = m/2$.

(b) *There is an* A^- *such that* $AA^- = A^-A$ *if and only if* $r(A^2) = r(A)$.

(c) *The rank of* $AA^- - A^-A$ *is invariant with respect to the choice of* A^- *if and only if* $A^2 = 0$ *or* $r(A^2) = 2r(A) - m$.

The two rank formulas (21.69) and (21.70) manifest that the maximal and minimal ranks of $AA^- - A^-A$ are even. Recall from (6.1) that the rank of $AA^\dagger - A^\dagger A$ is even, too. Thus we have the following conjecture.

Conjecture 21.20. *Let* $A \in \mathcal{F}^{m\times m}$ *be given. Then the rank of the matrix expression* $AA^- - A^-A$ *is even for any* A^-.

In the same way we can establish the following. The details are omitted.

Theorem 21.21. *Let* $A \in \mathcal{F}^{m\times k}$ *and* $B \in \mathcal{F}^{l\times m}$ *be given. Then*

(a) *The maximal and the minimal ranks of* $AA^- - B^-B$ *with respect to* A^- *and* B^- *are*

$$\max_{A^-,\, B^-} r(\, AA^- - B^-B\,) = \min\{\, 2m - r(A) - r(B), \quad r(A) + r(B)\,\}, \tag{21.73}$$

$$\min_{A^-,\, B^-} r(\, AA^- - B^-B\,) = r(A) + r(B) - 2r(BA). \tag{21.74}$$

(b) *There are A^- and B^- such that $AA^- - B^-B$ is nonsingular if and only if $r(A) + r(B) = m$.*

(c) *There are A^- and B^- such that $AA^- = B^-B$ if and only if $r(A) + r(B) = 2r(BA)$.*

(d) *The rank of $AA^- - B^-B$ is invariant with respect to the choice of A^- and B^- if and only if $BA = 0$ or $r(BA) = r(A) + r(B) - m$.*

As for extreme ranks of $AA^- + B^-B$ we shall present them in Chapter 27. Moreover, we can also determine the maximal and the minimal ranks of $BB^-A - AC^-C$.

Theorem 21.22. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ and $C \in \mathcal{F}^{l\times n}$ be given. Then*

$$\max_{B^-,\,C^-} r(\,BB^-A - AC^-C\,)$$
$$= \min\left\{ r[\,A,\ B\,], \quad r\begin{bmatrix} A \\ C \end{bmatrix}, \quad r(B) + r(C), \quad r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,] - r(B) - r(C) \right\}, \tag{21.75}$$

and

$$\min_{B^-,\,C^-} r(\,BB^-A - AC^-C\,) = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,] + r(B) + r(C) - 2r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}. \tag{21.76}$$

Proof. According to (4.1), the rank of $BB^-A - AC^-C$ can be written as

$$\begin{aligned} r(\,BB^-A - AC^-C\,) &= r\begin{bmatrix} BB^-A \\ C^-C \end{bmatrix} + r[\,AC^-C,\ BB^-\,] - r(BB^-) - r(C^-C) \\ &= r\begin{bmatrix} BB^-A \\ C \end{bmatrix} + r[\,AC^-C,\ B\,] - r(B) - r(C). \end{aligned}$$

Hence

$$\max_{B^-,\,C^-} r(\,BB^-A - AC^-C\,) = \max_{B^-} r\begin{bmatrix} BB^-A \\ C \end{bmatrix} + \max_{C^-} r[\,AC^-C,\ B\,] - r(B) - r(C), \tag{21.77}$$

$$\min_{B^-,\,C^-} r(\,BB^-A - AC^-C\,) = \min_{B^-} r\begin{bmatrix} BB^-A \\ C \end{bmatrix} + \min_{C^-} r[\,AC^-C,\ B\,] - r(B) - r(C). \tag{21.78}$$

According to (21.4) and (21.5), we easily find that

$$\begin{aligned} &\max_{B^-} r\begin{bmatrix} BB^-A \\ C \end{bmatrix} = \min\left\{ r(B) + r(C), \quad r\begin{bmatrix} A \\ C \end{bmatrix} \right\}, \\ &\min_{B^-} r\begin{bmatrix} BB^-A \\ C \end{bmatrix} = r(B) + r(C) + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}, \\ &\max_{C^-} r[\,AC^-C,\ B\,] = \min\{\, r[\,A,\ B\,], \quad r(B) + r(C)\,\}, \\ &\min_{C^-} r[\,AC^-C,\ B\,] = r(B) + r(C) + r[\,A,\ B\,] - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}. \end{aligned}$$

Putting them in (21.77) and (21.78) yields (21.75) and (21.76). $\square$

Corollary 21.23. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ and $C \in \mathcal{F}^{l\times m}$ be given.*

(a) *Assume A is square. Then there are B^- and C^- such that $BB^-A - AC^-C$ is nonsingular if and only if A, B and C satisfy the following rank equality*

$$r\begin{bmatrix} A \\ C \end{bmatrix} = r[\,A,\ B\,] = r(B) + r(C) = m. \tag{21.79}$$

(b) *There are B^- and C^- such that $BB^-A = AC^-C$ if and only if A, B and C satisfy the rank additivity condition*

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r(B) = r[\,A,\ B\,] + r(C). \tag{21.80}$$

(c) *The rank of* $BB^-A - AC^-C$ *is invariant with respect to the choice of* B^- *and* C^- *if and only if*

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(B) + r(C), \quad \text{or} \quad r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} = r[\, A, \ B\,]. \tag{21.81}$$

Proof. Follows from (21.75) and (21.76). □

Theorem 21.24. *Let* $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{k\times m}$ *and* $C \in \mathcal{F}^{n\times l}$ *be given.*

(a) *The maximal and minimal ranks of* $B^-BA - ACC^-$ *with respect to* B^- *and* C^- *are given by*

$$\max_{B^-,\, C^-} r(\, B^-BA - ACC^-\,) = \min\{\, r(BA) + r(AC), \quad m + n - r(B) - r(C)\,\}, \tag{21.82}$$

$$\min_{B^-,\, C^-} r(\, B^-BA - ACC^-\,) = r(BA) + r(AC) - 2r(BAC). \tag{21.83}$$

(b) *There are* B^- *and* C^- *such that* $B^-BA = ACC^-$ *if and only if* A, B *and* C *satisfy the rank equality* $r(BAC) = r(BA) = r(AC)$.

(c) *The rank of* $B^-BA - ACC^-$ *is invariant with respect to the choice of* B^- *and* C^- *if and only if*

$$BAC = 0 \quad \text{or} \quad r(BAC) = r(B) + r(AC) - m = r(BA) + r(C) - n. \tag{21.84}$$

The proof of Theorem 21.24 is similar to that of Theorem 21.22 and is, therefore, omitted. Replacing A in (21.82) and (21.83) by A^{k-1}, and B and C in (21.82) and (21.83) by A, we directly obtain the following.

Corollary 21.25. *Let* $A \in \mathcal{F}^{m\times m}$ *be given.*

(a) *The maximal and the minimal ranks of* $A^kA^- - A^-A^k$ *with respect to* A^- *are*

$$\max_{A^-} r(\, A^kA^- - A^-A^k\,) = \min\{\, 2m - 2r(A), \quad 2r(A^k)\,\}, \tag{21.85}$$

$$\min_{A^-} r(\, A^kA^- - A^-A^k\,) = 2r(A^k) - 2r(A^{k+1}). \tag{21.86}$$

(b) *There is an* A^- *such that* $A^kA^- - A^-A^k$ *is nonsingular if and only if* m *is even and* $r(A^k) = r(A) = m/2$.

(c) *There is an* A^- *such that* $A^kA^- = A^-A^k$ *if and only if* $r(A^{k+1}) = r(A^k)$.

(d) *The rank of* $A^kA^- - A^-A^k$ *is invariant with respect to the choice of* A^- *if and only if* $A^{k+1} = 0$ *or* $r(A^{k+1}) = r(A^k) + r(A) - m$.

Theorem 21.26. *Let* $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times p}$ *and* $C \in \mathcal{F}^{l\times n}$ *be given. Then*

(a) *The maximal and minimal ranks of* $I_m - AA^- - BB^-$ *and* $I_m - A^-A - C^-C$ *with respect to* A^-, B^- *and* C^- *are*

$$\max_{A^-, B^-} r(\, I_m - AA^- - BB^-\,) = m - |\, r(A) - r(B)\,|, \tag{21.87}$$

$$\min_{A^-, B^-} r(\, I_m - AA^- - BB^-\,) = m + r(A) + r(B) - 2r[\, A, \ B\,]. \tag{21.88}$$

$$\max_{A^-, C^-} r(\, I_n - A^-A - C^-C\,) = n - |\, r(A) - r(C)\,|, \tag{21.89}$$

$$\min_{A^-, B^-} r(\, I_n - A^-A - C^-C\,) = n + r(A) + r(C) - 2r\begin{bmatrix} A \\ C \end{bmatrix}. \tag{21.90}$$

(b) *There are* A^- *and* B^- *such that* $I_m - AA^- - BB^-$ *is nonsingular if and only if* $r(A) = r(B)$.

(c) *There are* A^- *and* B^- *such that* $AA^- + BB^- = I_m$ *if and only if* $r[\, A, \ B\,] = r(A) + r(B) = m$.

(d) *There are* A^- *and* C^- *such that* $I_n - A^-A - C^-C$ *is nonsingular if and only if* $r(A) = r(C)$.

(e) *There are* A^- *and* C^- *such that* $A^-A + C^-C = I_n$ *if and only if* $r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r(C) = n$.

Proof. According to (3.8), the rank of $I_m - AA^- - BB^-$ can be written as

$$\begin{aligned} r(\, I_m - AA^- - BB^-\,) &= r(AA^-BB^-) + r(BB^-AA^-) - r(AA^-) - r(BB^-) + m \\ &= r(AA^-B) + r(BB^-A) - r(A) - r(B) + m. \end{aligned}$$

Hence

$$\max_{A^-,\, B^-} r(\, I_m - AA^- - BB^- \,) = \max_{A^-} r(AA^-B) + \max_{B^-} r(BB^-A) - r(A) - r(B) + m, \tag{21.91}$$

$$\min_{A^-,\, B^-} r(\, I_m - AA^- - BB^- \,) = \min_{A^-} r(AA^-B) + \min_{B^-} r(BB^-A) - r(A) - r(B) + m. \tag{21.92}$$

It follows from Corollaries 21.7 and 21.9 that

$$\max_{A^-} r(AA^-B) = \max_{B^-} r(BB^-A) = \min\{\, r(A), \quad r(B)\, \},$$
$$\min_{A^-} r(AA^-B) = \min_{B^-} r(BB^-A) = r(A) + r(B) - r[\, A, \;\, B\,].$$

Putting them in (21.91) and (21.92) yields (21.87) and (21.88). By the similar approach, we can get (21.88) and (21.89). □

Replace B in (21.88) and (21.90) by $I_m - A$, we get by noticing (1.11)

$$\begin{aligned} &\min_{A^-,(I_m-A)^-} r[\, I_m - AA^- - (I_m - A)(I_m - A)^- \,] \\ &= \min_{A^-,(I_m-A)^-} r[\, I_m - A^-A - (I_m - A)^-(I_m - A)\,] \\ &= r(\, A - A^2 \,). \end{aligned} \tag{21.93}$$

Thus the following three statements are equivalent:

(a) There are A^- and $(I_m - A)^-$ such that $AA^- + (I_m - A)(I_m - A)^- = I_m$.

(b) There are A^- and $(I_m - A)^-$ such that $A^-A + (I_m - A)^-(I_m - A) = I_m$.

(c) A is idempotent.

The two two statements in (a) and (b) could be regarded as new characterizations of idempotent matrix.

Replace A and B in (21.88) and (21.90) by $I_m + A$ and $I_m - A$, respectively, we get by noticing (1.12)

$$\begin{aligned} &\min_{(I_m+A)^-,(I_m-A)^-} r[\, I_m - (I_m + A)(I_m + A)^- - (I_m - A)(I_m - A)^- \,] \\ &= \min_{(I_m+A)^-,(I_m-A)^-} r[\, I_m - (I_m + A)^-(I_m + A) - (I_m - A)^-(I_m - A)\,] \\ &= r(\, I_m - A^2 \,). \end{aligned} \tag{21.94}$$

Thus the following three statements are equivalent:

(a) There are $(I_m + A)^-$ and $(I_m - A)^-$ such that $(I_m + A)(I_m + A)^- + (I_m - A)(I_m - A)^- = I_m$.

(b) There are $(I_m + A)^-$ and $(I_m - A)^-$ such that $(I_m + A)^-(I_m + A) + (I_m - A)^-(I_m - A) = I_m$.

(c) A is involutory.

The two two statements in (a) and (b) could be regarded as new characterizations of involutory matrix.

In general, suppose that $p(x)$ and $q(x)$ are two polynomials without common roots. Then there is by (21.88), (21.90) and (1.17) the following

$$\min_{p^-(A),\, q^-(A)} r[\, I_m - p(A)p^-(A) - q(A)q^-(A)\,] = \min_{p^-(A),\, q^-(A)} r[\, I_m - p^-(A)p(A) - q^-(A)q(A)\,] = r[\, p(A)q(A)\,]. \tag{21.95}$$

Thus there exist $p^-(x)$ and $q^-(x)$ such that $p(A)p^-(A) + q(A)q^-(A) = I_m$ if and only if $p(A)q(A) = 0$. We leave its verification to the reader.

Theorem 21.27. *Let $A \in \mathcal{F}^{m\times m}$ be given. Then*

$$\max_{A^-} r(\, I_m \pm A^k - AA^- \,) = \min_{A^-} r(\, I_m \pm A^k - AA^- \,) = r(A^{k+1}) - r(A) + m, \tag{21.96}$$

$$\max_{A^-} r(\, I_m \pm A^k - A^-A \,) = \min_{A^-} r(\, I_m \pm A^k - A^-A \,) = r(A^{k+1}) - r(A) + m. \tag{21.97}$$

That is, the equalities

$$r(A^{k+1}) = r(A) - m + r(\, I_m \pm A^k - AA^- \,) = r(A) - m + r(\, I_m \pm A^k - A^-A \,) \tag{21.98}$$

hold for any A^-.

Proof. Applying (21.4) and (21.5) to $I_m \pm A^k - AA^-$ and $I_m \pm A^k - A^-A$ yields the desired results. $\square$

We leave the proof of the following result to the reader.

Theorem 21.28. *Let $A \in \mathcal{F}^{m\times m}$ be given. Then*

$$\max_{(I_m - A)^-} r\left[(I_m - A)^- - \sum_{i=0}^{k-1} A^i\right] = \min\{\, m, \quad m + r(\, I_m - A^k\,) - r(A)\,\}, \tag{21.99}$$

and

$$\begin{aligned} &\min_{(I_m - A)^-} r\left[(I_m - A)^- - \sum_{i=0}^{k-1} A^i\right] \\ &= \min_{(I_m - A)^-} r[\,(I_m - A)(I_m - A)^- - (\,I_m - A^k\,)\,] \\ &= \min_{(I_m - A)^-} r[\,(I_m - A)^-(I_m - A) - (\,I_m - A^k\,)\,] = r(\,A^k - A^{k+1}\,). \end{aligned} \tag{21.100}$$

In particular, the following four statements are equivalent:

(a) $\sum_{i=0}^{k-1} A^i \in \{(I_m - A)^-\}$.

(b) $(I_m - A)(I_m - A)^- = I_m - A^k$.

(c) $(I_m - A)^-(I_m - A) = I_m - A^k$.

(d) $A^{k+1} = A^k$, *i.e., A is quasi-idempotent.*

A parallel result to (21.100) is

$$\min_{(\sum_{i=0}^{k-1} A^i)^-} r\left[\left(\sum_{i=0}^{k-1} A^i\right)^- - (I_m - A)\right] = r(\,A^k + A^{k+1} + \cdots + A^{2k}\,). \tag{21.101}$$

It implies that $I_m - A \in \{\,(\sum_{i=0}^{k-1} A^i)^-\}$ if and only if $A^k + A^{k+1} + \cdots + A^{2k} = 0$.

It is expected that one can further establish numerous rank equalities among A^-, $(I - A)^-$ and polynomials of A, and then derive from them various conclusions related A^- and $(I - A)^-$. we leave this work to the reader.

In the remainder of this chapter, we consider the rank of the difference $A^- - PN^-Q$ and then present some of their consequences.

Theorem 21.28. *Let $A \in \mathcal{F}^{m\times n}$, $P \in \mathcal{F}^{n\times p}$ $N \in \mathcal{F}^{q\times p}$ and $Q \in \mathcal{F}^{q\times n}$ be given. Then*

$$\min_{A^-,\, N^-} r(\,A^- - PN^-Q\,) = r(\,N - QAP\,) + r(A) + r(N) - r[\,A,\ QAP\,] - r\begin{bmatrix} A \\ QAP \end{bmatrix}. \tag{21.102}$$

In particular, there are A^- and N^- such that $A^- = PN^-Q$ if and only if

$$r(\,N - QAP\,) = r\begin{bmatrix} N \\ QAP \end{bmatrix} + r[\,N,\ QAP\,] - r(A) - r(N). \tag{21.103}$$

Proof. From (21.5) we can get the following simple result

$$\min_{A^-} r(\,A^- - D\,) = r(\,A - ADA\,).$$

Applying it to $A^- - PN^-Q$ we first get

$$\min_{A^-} r(\,A^- - PN^-Q\,) = r(\,A - APN^-QA\,).$$

Next applying (21.5) to its right hand side and simplifying, we then have

$$\min_{N^-} r(\, A - APN^-QA\,) = r(\, N - QAP\,) + r(A) + r(N) - r[\,A,\ QAP\,] - r\begin{bmatrix} A \\ QAP \end{bmatrix}.$$

Combining the above two equalities results in (21.102). □

Clearly (21.29) could be regarded as a special case of (21.102). Now applying (21.102) to the block matrix $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$, we can simply get following

$$\min_{A^-,\,M^-} r\left(A^- - [\,I_n,\ 0]M^-\begin{bmatrix} I_m \\ 0 \end{bmatrix}\right) = r(A) + r(M) + r\begin{bmatrix} 0 & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix}. \tag{21.104}$$

This result implies that there is an M^- which upper left block is an inner inverse of A if and only if

$$r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} + r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} = r(A) + r(M) + r\begin{bmatrix} 0 & B \\ C & D \end{bmatrix}. \tag{21.105}$$

Applying (21.102) to a block circulant matrix M generated by k matrices $A_1,\ A_2,\ \cdots,\ A_k$ and their sum $A = A_1 + A_2 + \cdots + A_k$, we can also get

$$\min_{A^-,\,M^-} r\left(A^- - \frac{1}{k}[\,I,\ \cdots\ I\,]M^-\begin{bmatrix} I \\ \vdots \\ I \end{bmatrix}\right) = 0, \tag{21.106}$$

that is, there must exist A^- and M^- such that

$$A^- = [\,I,\ \cdots,\ I\,]M^-\begin{bmatrix} I \\ \vdots \\ I \end{bmatrix}. \tag{21.107}$$

Besides the result in Theorem 21.28, we can also determine the relationship between the two matrix sets $\{PN^-Q\}$ and $\{A^-\}$. Here we only list the main results without proofs.

Theorem 21.29. *Let $A \in \mathcal{F}^{m\times n}$, $N \in \mathcal{F}^{k\times l}$, $P \in \mathcal{F}^{n\times l}$ and $Q \in \mathcal{F}^{k\times m}$ be given with $r(P) = n$ and $r(A) = m$. Then*

$$\max_{N^-} r(\, A - APN^-QA\,) = \min\{\, r(A),\quad r(N - QAP) + r(A) - r(N)\,\}, \tag{21.108}$$

In particular, the following set inclusion

$$\{PN^-Q\} \subseteq \{A^-\}, \tag{21.109}$$

holds if and only if

$$r(N - QAP) = r(N) - r(A). \tag{21.110}$$

Theorem 21.30. *Let $A \in \mathcal{F}^{m\times n}$, $N \in \mathcal{F}^{k\times l}$, $P \in \mathcal{F}^{n\times l}$ and $Q \in \mathcal{F}^{k\times m}$ be given with $r(P) = n$ and $r(A) = m$. Then*

$$\max_{A^-}\min_{N^-} r(\, A^- - PN^-Q\,) = \min\left\{ n + r(N) - r\begin{bmatrix} N \\ P \end{bmatrix},\quad m + r(N) + r[\,N,\ Q\,],\right.$$
$$\left. m + n + r(N - QAP) + r(N) - r(A) - r[\,N,\ Q\,] - r\begin{bmatrix} N \\ P \end{bmatrix}\right\}. \tag{21.111}$$

In particular, the following set inclusion

$$\{A^-\} \subseteq \{PN^-Q\}, \tag{21.112}$$

holds if and only if $R(N) \cap R(Q) \neq \{0\}$, *or* $R(N^T) \cap R(P^T) \neq \{0\}$, *or*

$$r(N - QAP) = r \begin{bmatrix} N \\ P \end{bmatrix} + r[\, N,\, Q\,] - r(N) + r(A) - m - n. \tag{21.113}$$

Theorem 21.31. *Let* $A \in \mathcal{F}^{m \times n}$, $N \in \mathcal{F}^{k \times l}$, $P \in \mathcal{F}^{n \times l}$ *and* $Q \in \mathcal{F}^{k \times m}$ *be given with* $r(P) = n$ *and* $r(A) = m$. *Under the condition* $R(N) \cap R(Q) = \{0\}$ *and* $R(N^T) \cap R(P^T) = \{0\}$, *the following equality*

$$\{PN^-Q\} = \{A^-\} \tag{21.114}$$

holds if and only if

$$r \begin{bmatrix} N & Q \\ P & 0 \end{bmatrix} = r \begin{bmatrix} N \\ P \end{bmatrix} + r(Q) = r[\, N,\, Q\,] + r(P), \tag{21.115}$$

and

$$A = -[\, 0,\ I_m\,] \begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}^- \begin{bmatrix} 0 \\ I_n \end{bmatrix}. \tag{21.116}$$

Based on the above three theorems, we can establish the following several results.

Theorem 21.32. *Let* $A_1, A_2, \cdots, A_k$ *be given matrices of the same size, and let* $A = A_1 + A_2 + \cdots + A_k$. *Denote by* M *the block circulant matrix generated by* $A_1, A_2, \cdots, A_k$. *Then* A *and* M *satisfy*

$$\{\, A^- \,\} = \left\{ \frac{1}{k}[\, I, \cdots, I\,]M^- \begin{bmatrix} I \\ \vdots \\ I \end{bmatrix} \right\}. \tag{21.117}$$

Theorem 21.33. *Let* $A + iB$ *be an* $m \times n$ *complex matrix. Then*

$$\{\, (A + iB)^- \,\} = \left\{ \frac{1}{2}[\, I_n,\, iI_n\,] \begin{bmatrix} A & -B \\ B & A \end{bmatrix}^- \begin{bmatrix} I_m \\ -iI_m \end{bmatrix} \right\}. \tag{21.118}$$

Theorem 13.34. *Let* $A_0 + iA_1 + iA_2 + kA_3$ *be an* $m \times n$ *real quaternion matrix. Then*

$$\left\{\, (\, A_0 + iA_1 + iA_2 + kA_3\,)^- \,\right\} = \left\{ \frac{1}{4}[\, I_n,\ iI_n,\ jI_n,\ kI_n\,] \begin{bmatrix} A_0 & -A_1 & -A_2 & -A_3 \\ A_1 & A_0 & A_3 & -A_2 \\ A_2 & -A_3 & A_0 & A_1 \\ A_3 & A_2 & -A_1 & A_0 \end{bmatrix}^- \begin{bmatrix} I_m \\ -iI_m \\ -jI_m \\ -kI_m \end{bmatrix} \right\}. \tag{21.119}$$

Chapter 22

Generalized inverses of multiple matrix products

Generalized inverses of products of matrices have been an attractive topic in the theory of generalized inverses matrices. Various results related to reverse order laws for g-inverses, reflexive g-inverses, and the Moore-Penrose inverses of matrix products can be found in the literature. Generally speaking, this work has two main directions according the classification of generalized inverses of matrices, one of which is concerned with reverse order laws for the Moore-Penrose inverses of matrix products. Up till now, necessary and sufficient conditions for the reverse order laws $(AB)^{\dagger} = B^{\dagger}A^{\dagger}$, $(ABC)^{\dagger} = (BC)^{\dagger}B(AB)^{\dagger}$, $(ABC)^{\dagger} = C^{\dagger}B^{\dagger}A^{\dagger}$, $(A_1A_2\cdots A_k)^{\dagger} = A_k^{\dagger}\cdots A_2^{\dagger}A_1^{\dagger}$ to hold have well been established. The other direction of this work is concerned with reverse order laws for inner inverses, reflexive inner inverses, as well as several other types of generalized inverses of matrix products. Some earlier and recent work gives a complete consideration for the reverse order laws $(AB)^- = B^-A^-$, $(AB)_r^- = B_r^-A_r^-$, $(AB)^{\dagger} = B_{mr}^-A_{lr}^-$, and $(AB)_{MN}^{\dagger} = B^-A^-$ (see, e.g., [12, 41, 118, 123, 149, 150, 151]). In this chapter we first present some rank equalities related to $(BC)^-B(AB)^-$ and $C^-B^-A^-$, and then apply them to establish the relationships between $(BC)^-B(AB)^-$ and $(ABC)^-$, $C^-B^-A^-$ and $(ABC)^-$.

The following lemma comes directly from (22.4) and (22.5), which will be used in the sequel.

Lemma 22.1. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$, $C \in \mathcal{F}^{l\times n}$ and $D \in \mathcal{F}^{l\times k}$ be given.*

(a) *If $R(D) \subseteq R(C)$ and $R(D^T) \subseteq R(B^T)$, then*

$$\max_{A^-} r(\, D - CA^-B\,) = \min\left\{ r(B), \quad r(C), \quad r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r(A) \right\}, \tag{22.1}$$

$$\min_{A^-} r(\, D - CA^-B\,) = r(A) - r[\,A,\ B\,] - r\begin{bmatrix} A \\ C \end{bmatrix} + r\begin{bmatrix} A & B \\ C & D \end{bmatrix}. \tag{22.2}$$

(b) *In particular,*

$$\max_{A^-} r(\, D - CA^-\,) = \min\{\, m, \quad r[\,C,\ D\,], \quad r(\, C - DA\,) - r(A) + m\,\}, \tag{22.3}$$

$$\max_{A^-} r(\, D - A^-B\,) = \min\left\{ n, \quad r\begin{bmatrix} B \\ D \end{bmatrix}, \quad r(\, B - AD\,) - r(A) + n \right\}. \tag{22.4}$$

Proof. Follows immediately from (21.5) and (21.6). □

It is quite obvious that there are $(AB)^- \in \{\,(AB)^-\}$ and $(BC)^- \in \{\,(BC)^-\}$ such that $(BC)^-B(AB)^- \subseteq \{\,(ABC)^-\}$ if and only if

$$\min_{(AB)^-,\,(BC)^-} r[\,ABC - (ABC)(BC)^-B(AB)^-(ABC)\,] = 0, \tag{22.5}$$

and the set inclusion $\{(BC)^-B(AB)^-\} \subseteq \{(ABC)^-\}$ holds if and only if

$$\max_{(AB)^-,\,(BC)^-} r[\,ABC - (ABC)(BC)^-B(AB)^-(ABC)\,] = 0. \tag{22.6}$$

These two equivalence statements clearly show that the relationship between $(BC)^{-}B(AB)^{-}$ and $(ABC)^{-}$ can be characterized by extreme ranks of matrix expressions involving $(BC)^{-}, (AB)^{-}$, and $(ABC)^{-}$. According to the rank formulas in Lemma 22.1, we easily find the following special result.

Theorem 22.2. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{n\times p}$, and $C \in \mathcal{F}^{p\times q}$ be given, and let $M = ABC$. Then*

$$\min_{(AB)^{-}} r[\, M - M(BC)^{-}B(AB)^{-}M\,] = 0, \qquad \min_{(BC)^{-}} r[\, M - M(BC)^{-}B(AB)^{-}M\,] = 0, \tag{22.7}$$

and

$$\max_{(AB)^{-},\,(BC)^{-}} r[\, M - M(BC)^{-}B(AB)^{-}M\,] = \min\{\, r(M), \quad r(M) - r(AB) - r(BC) + r(B)\,\}\,. \tag{22.8}$$

Proof. Applying (22.2) to $M - M(BC)^{-}B(AB)^{-}M$, we find

$$\begin{aligned} &\min_{(AB)^{-}} r[\, M - M(BC)^{-}B(AB)^{-}M\,] \\ &\quad = r(AB) - r[\, AB,\ M\,] - r\begin{bmatrix} AB \\ M(BC)^{-}B \end{bmatrix} + r\begin{bmatrix} AB & M \\ M(BC)^{-}B & M \end{bmatrix} \\ &\quad = r(AB) - r[\, AB,\ 0\,] - r\begin{bmatrix} AB \\ M(BC)^{-}B \end{bmatrix} + r\begin{bmatrix} AB & 0 \\ M(BC)^{-}B & 0 \end{bmatrix} = 0, \end{aligned}$$

$$\begin{aligned} &\min_{(BC)^{-}} r[\, M - M(BC)^{-}B(AB)^{-}M\,] \\ &\quad = r(BC) - r[\, BC,\ B(AB)^{-}M\,] - r\begin{bmatrix} BC \\ M \end{bmatrix} + r\begin{bmatrix} BC & B(AB)^{-}M \\ M & M \end{bmatrix} \\ &\quad = r(BC) - r[\, BC,\ B(AB)^{-}M\,] - r\begin{bmatrix} BC \\ 0 \end{bmatrix} + r\begin{bmatrix} BC & B(AB)^{-}M \\ 0 & 0 \end{bmatrix} = 0. \end{aligned}$$

Both of them are (22.7). Next by (22.1), we find

$$\begin{aligned} \max_{(AB)^{-}} r[\, M - M(BC)^{-}B(AB)^{-}M\,] &= \min\left\{ r[M(BC)^{-}B], \quad r(M), \quad r\begin{bmatrix} AB & M \\ M(BC)^{-}B & M \end{bmatrix} - r(AB) \right\} \\ &= \min\left\{ r(M), \quad r\begin{bmatrix} AB \\ M(BC)^{-}B \end{bmatrix} - r(AB) \right\}. \end{aligned}$$

Accroding to (21.4) and $r(M) \le r(BC)$, we also find that

$$\begin{aligned} \max_{(BC)^{-}} r\begin{bmatrix} AB \\ M(BC)^{-}B \end{bmatrix} &= \max_{(BC)^{-}} r\left(\begin{bmatrix} AB \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ -M \end{bmatrix} (BC)^{-}B \right) \\ &= \min r \left\{ r\begin{bmatrix} AB & 0 \\ 0 & -M \end{bmatrix}, \quad r\begin{bmatrix} AB \\ 0 \\ B \end{bmatrix}, \quad r\begin{bmatrix} BC & B \\ 0 & AB \\ -M & 0 \end{bmatrix} - r(BC) \right\} \\ &= \min\{\, r(AB) + r(M), \quad r(B), \quad r(B) + r(M) - r(BC)\,\} \\ &= \min\{\, r(AB) + r(M), \quad r(B) + r(M) - r(BC)\,\}. \end{aligned}$$

Combining the above two equalities, we obtain

$$\begin{aligned} \max_{(AB)^{-},\,(BC)^{-}} r[\, M - M(BC)^{-}B(AB)^{-}M\,] &= \min\left\{ r(M), \quad \max_{(BC)^{-}} r\begin{bmatrix} AB \\ M(BC)^{-}B \end{bmatrix} - r(AB) \right\} \\ &= \min\{\, r(M), \quad r(M) - r(AB) - r(BC) + r(B)\,\}, \end{aligned}$$

which is exactly (22.8). $\square$.

Combining (22.5) and (22.6) with (22.7) and (22.8), we obtain the main result in this chapter.

Theorem 22.3. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{n\times p}$, and $C \in \mathcal{F}^{p\times q}$ be given.*

(a) *For every* $(AB)^- \in \{(AB)^-\}$, *there must be a* $(BC)^- \in \{(BC)^-\}$ *such that* $(BC)^-B(AB)^- \in \{(ABC)^-\}$ *holds.*

(b) *For every* $(BC)^- \in \{(BC)^-\}$, *there must be an* $(AB)^- \in \{(AB)^-\}$ *such that* $(BC)^-B(AB)^- \in \{(ABC)^-\}$ *holds.*

(c) *The set inclusion* $\{(BC)^-B(AB)^-\} \subseteq \{(ABC)^-\}$ *holds if and only if*

$$ABC = 0 \quad or \quad r(ABC) = r(AB) + r(BC) - r(B). \tag{22.9}$$

(d) *In particular, if* $r(ABC) = r(B)$, *then* $\{(BC)^-B(AB)^-\} \subseteq \{(ABC)^-\}$ *holds.*

As a direct consequence by setting $B = I$ in Theorem 22.3, we obtain the following.

Corollary 22.4. *Let* $A \in \mathcal{F}^{m\times n}$ *and* $B \in \mathcal{F}^{n\times p}$ *be given.*

(a) *For every* $A^- \in \{A^-\}$, *there must be a* $B^- \in \{B^-\}$ *such that* $B^-A^- \in \{(AB)^-\}$ *holds.*

(b) *For every* $B^- \in \{B^-\}$, *there must be an* $A^- \in \{A^-\}$ *such that* $B^-A^- \in \{(AB)^-\}$ *holds.*

(c) *The set inclusion* $\{B^-A^-\} \subseteq \{(AB)^-\}$ *holds if and only if*

$$AB = 0 \quad or \quad r(AB) = r(A) + r(B) - n. \tag{22.10}$$

Necessary and sufficient conditions for $\{B^-A^-\} \subseteq \{(AB)^-\}$ were previously examined by Gross in [54], Werner in [149] and [150]. The results given there are in fact equivalent to (22.10) The results in Theorem 22.3 and Corollary 22.4 can help us to establish various relationship between generalized inverses of matrices. We next present one of them.

Corollary 22.5. *Let* $A,\, B \in \mathcal{F}^{m\times n}$ *be given and let* $M = \mathrm{diag}(A,\, B)$, $N = A + B$, $S = \begin{bmatrix} A \\ B \end{bmatrix}$, *and* $T = [\,A,\ B\,]$. *Then*

$$\max_{S^-,\,T^-} r\left(N - N\begin{bmatrix} A \\ B \end{bmatrix}^- M[\,A,\ B\,]^-N\right) = \min\left\{r(N), \quad r(N) - r\begin{bmatrix} A \\ B \end{bmatrix} - r[\,A,\ B\,] + r(A) + r(B)\right\}. \tag{22.11}$$

In particular, the set inclusion

$$\left\{\begin{bmatrix} A \\ B \end{bmatrix}^- \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}[\,A,\ B\,]^-\right\} \subseteq \{\,(\,A + B\,)^-\,\} \tag{22.12}$$

holds if and only if

$$A + B = 0 \quad or \quad r(\,A + B\,) = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\,A,\ B\,] - r(A) - r(B). \tag{22.13}$$

Proof. Writing $A + B$ as the product

$$A + B = [\,I_m,\ I_m\,]\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} I_n \\ I_n \end{bmatrix} = PDQ.$$

Then (22.11) follows from (22.8). Consequently (22.12) and (22.13) follow form (22.11). □

Eq. (22.12) has some interesting consequences. For example, in the case of $N = A + (I_m - A) = I_m$, then

$$\max_{S^-,\,T^-} r\left(I_m - \begin{bmatrix} A \\ I_m - A \end{bmatrix}^- \begin{bmatrix} A & 0 \\ 0 & I_m - A \end{bmatrix}[\,A,\ I_m - A\,]^-\right) = r(A) + r(I_m - A) - m = r(A - A^2).$$

This implies that A is idempotent if and only if

$$\begin{bmatrix} A \\ I_m - A \end{bmatrix}^- \begin{bmatrix} A & 0 \\ 0 & I_m - A \end{bmatrix}[\,A,\ I_m - A\,]^- = I_m$$

holds for any inner inverses in it. This fact could be regarded as a characterization for idempotent matrix. In addition for any two idempotent matrices A and B, there is

$$\max_{S^-,\,T^-} r\left((A-B) - (A-B)\begin{bmatrix} A \\ -B \end{bmatrix}^- \begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix} [\,A,\ -B\,]^-(A-B) \right) = 0.$$

Thus

$$\left\{ \begin{bmatrix} A \\ -B \end{bmatrix}^- \begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix} [\,A,\ -B\,]^- \right\} \subseteq \{\,(\,A-B\,)^-\,\}$$

holds for any two idempotent matrices A and B.

An extension of Corollary 22.5 is given below.

Corollary 22.6. *Let $A_1,\, A_2,\, \cdots,\, A_k \in \mathcal{F}^{m\times n}$ be given. The set inclusion*

$$\left\{ \begin{bmatrix} A_1 \\ \vdots \\ A_k \end{bmatrix}^- \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix} [\,A_1,\, \cdots,\, A_k\,]^- \right\} \subseteq \{\ (\,A_1 + \cdots + A_k\,)^-\ \}$$

holds if and only if

$$A_1 + \cdots + A_k = 0 \quad \text{or} \quad r(\,A_1 + \cdots + A_k\,) = r\begin{bmatrix} A_1 \\ \vdots \\ A_k \end{bmatrix} + r[\,A_1,\, \cdots,\, A_k\,] - r(A_1) - \cdots - r(A_k).$$

For some products of matrices, the two equalities in (22.10) and (22.11) are satisfied, for example, (1.10)—(1.12), (1.14), (3.1), (4.1) and so on could be regarded as the special cases of (22.10) and (22.11). Thus based on them and Corollary 22.4(c) and Corollary 22.5(c), one can establish various set inclusions for inner inverses of products of matrices. We leave them to the reader.

Without much effort, we can also find a necessary and sufficient condition for the set inclusion $\{\,C^-B^-A^-\}$ $\subseteq \{\,(ABC)^-\}$ to hold. A rank formula related to this inclusion can be established using the rank formulas in (22.3) and (22.4).

Theorem 22.7. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{n\times p}$ and $C \in \mathcal{F}^{p\times q}$ be given, and let $M = ABC$. Then*

$$\max_{A^-,\,B^-,\,C^-} r(\,M - MC^-B^-A^-M\,) = \min\{\,r(M), \quad r(M) - r(A) - r(B) - r(C) + n + p\,\}. \tag{22.14}$$

Proof. We determine the maximal rank of $M - MC^-B^-A^-M$ subject to A^-, B^-, and C^- through the following step

$$\max_{A^-,\,B^-,\,C^-} r(\,M - MC^-B^-A^-M\,) = \max_{C^-}\max_{B^-}\max_{A^-} r(\,M - MC^-B^-A^-M\,). \tag{22.15}$$

According to (21.4) we first find

$$\begin{aligned}
&\max_{A^-} r(\,M - MC^-B^-A^-M\,) \\
&= \min\left\{ r[\,M,\ MC^-B^-\,], \quad r\begin{bmatrix} M \\ M \end{bmatrix}, \quad r\begin{bmatrix} A & M \\ MC^-B^- & M \end{bmatrix} - r(A) \right\} \\
&= \min\{\ r(M), \quad r(\,A - MC^-B^-\,) + r(M) - r(A)\,\}.
\end{aligned} \tag{22.16}$$

Next applying (22.3) to $A - MC^-B^-$ and noticing that $r(A) \le n$, we obtain

$$\begin{aligned}
\max_{B^-} r(\,A - MC^-B^-\,) &= \min\{\,r[\,A,\ MC^-\,], \quad r(\,MC^- - AB\,) - r(B) + n, \quad n\ \} \\
&= \min\{\ r(A), \quad r(\,MC^- - AB\,) - r(B) + n, \quad n\,\} \\
&= \min\{\ r(A), \quad r(\,AB - MC^-\,) - r(B) + n\ \}.
\end{aligned}$$

Consequently applying (22.3) to $AB - MC^-$ and the noticing that $r(AB) \le p$, we further find

$$\begin{aligned}\max_{C^-} r(\, AB - MC^- \,) &= \min\{\, r[\, AB,\ M\,], \quad r(\, M - ABC\,) - r(C) + p, \quad p\,\} \\ &= \min\{\, r(AB), \quad p - r(C)\,\}.\end{aligned}$$

Putting the above three results in (22.16) and noticing the Sylvester's law $r(AB) \ge r(A) + r(B) - n$, we eventually obtain

$$\begin{aligned}&\max_{A^-,\, B^-,\, C^-} r[\, M - MC^-B^-A^-M\,] \\ &= \min\left\{ r(M), \quad \max_{C^-}\max_{B^-} r(\, A - MC^-B^-\,) + r(M) - r(A) \right\} \\ &= \min\left\{ r(M), \quad \max_{C^-} r(\, AB - MC^-\,) + r(M) - r(A) - r(B) + n \right\} \\ &= \min\{\, r(M), \quad r(M) + r(AB) - r(A) - r(B) + n, \quad r(M) - r(A) - r(B) - r(C) + n + p\,\} \\ &= \min\{\, r(M), \quad r(M) - r(A) - r(B) - r(C) + n + p\,\},\end{aligned}$$

establishing (22.14). □

It is quite obvious that the set inclusion $\{\, C^-B^-A^-\} \subseteq \{\, (ABC)^-\}$ holds if and only if

$$\max_{A^-,\, B^-,\, C^-} r(\, M - MC^-B^-A^-M\,) = 0.$$

Thus from Theorem 3.1, we immediately obtain the following.

Theorem 22.8. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{n\times p}$, and $C \in \mathcal{F}^{p\times q}$ be given. Then the set inclusion $\{\, C^-B^-A^-\} \subseteq \{\, (ABC)^-\}$ holds if and only if*

$$ABC = 0 \quad or \quad r(ABC) = r(A) + r(B) + r(C) - n - p. \tag{22.17}$$

The results in Theorems 22.7 and 22.8 can easily be extended to inner inverse of multiple matrix products and its proof is omitted.

Theorem 22.9. *Let $A_1 \in \mathcal{F}^{n_1\times n_2}$, $A_2 \in \mathcal{F}^{n_2\times n_3}$, $\cdots$, $A_k \in \mathcal{F}^{n_k\times n_{k+1}}$ be given, and denote $M = A_1A_2\cdots A_k$. Then*

$$\max_{A_1^-,\, \cdots,\, A_k^-} r(\, M - MA_k^- \cdots A_1^- M\,) = \min\{\, r(M), \quad r(M) - r(A_1) - \cdots - r(A_k) + n_2 + \cdots + n_k\,\}. \tag{22.18}$$

Theorem 22.10. *Let $A_1 \in \mathcal{F}^{n_1\times n_2}$, $A_2 \in \mathcal{F}^{n_2\times n_3}$, $\cdots$, $A_k \in \mathcal{F}^{n_k\times n_{k+1}}$ be given. Then the set inclusion*

$$\{\, A_k^- \cdots A_2^- A_1^-\,\} \subseteq \{\, (\, A_1A_2\cdots A_k\,)^-\} \tag{22.19}$$

holds if and only if

$$A_1A_2\cdots A_k = 0 \quad or \quad r(\, A_1A_2\cdots A_k\,) = r(A_1) + r(A_2) + \cdots + r(A_k) - n_2 - n_3 - \cdots - n_k. \tag{22.20}$$

Combining Theorem 22.10 and the rank equality (1.16), we then get the following interesting result.

Theorem 22.11. *Let $A \in \mathcal{F}^{m\times m}$ be given, $\lambda_1, \lambda_2, \cdots, \lambda_k \in \mathcal{F}$ with $\lambda_i \ne \lambda_j$ for $i \ne j$, and denote $M = (\lambda_1 I - A)^{t_1}(\lambda_2 I - A)^{t_2}\cdots(\lambda_k I - A)^{t_k}$, where $t_1, t_2, \cdots, t_k$ are any positive integers. Then the following set inclusion holds*

$$\left\{\, [\, (\lambda_k I - A)^{t_k}]^- \cdots [\, (\lambda_2 I - A)^{t_2}]^- [\, (\lambda_1 I - A)^{t_1}\,]^- \,\right\} \subseteq \left\{\, M^-\,\right\}. \tag{22.21}$$

In general, the following set inclusion

$$\left\{\, [\, (\lambda_{i_1} I - A)^{t_{i_1}}]^- [\, (\lambda_{i_2} I - A)^{t_{i_2}}]^- \cdots [\, (\lambda_{i_k} I - A)^{t_{i_k}}\,]^- \,\right\} \subseteq \left\{\, M^-\,\right\}. \tag{22.22}$$

also holds, where $i_1,\, i_2,\, \cdots,\, i_k$ *are any permutation of* $1,\, 2,\, \cdots,\, k$.

As a special consequence, we see from (22.22) that the following six set inclusions all hold

$$\{\, A^-(I_m - A)^- \,\} \subseteq \{\, (A - A^2)^- \,\}, \qquad \{\, (I_m - A)^- A^- \,\} \subseteq \{\, (A - A^2)^- \,\}, \tag{22.23}$$

$$\{\, (I_m - A)^-(I_m + A)^- \,\} \subseteq \{\, (I_m - A^2)^- \,\}, \quad \{\, (I_m + A)^-(I_m - A)^- \,\} \subseteq \{\, (I_m - A^2)^- \,\}, \tag{22.24}$$

$$\{\, A^-(I_m - A)^-(I_m + A)^- \,\} \subseteq \{\, (A - A^3)^- \,\}, \quad \{\, A^-(I_m + A)^-(I_m - A)^- \,\} \subseteq \{\, (A - A^3)^- \,\}, \tag{22.25}$$

$$\{\, (I_m - A)^- A^-(I_m + A)^- \,\} \subseteq \{\, (A - A^3)^- \,\}, \quad \{\, (I_m + A)^- A^-(I_m - A)^- \,\} \subseteq \{\, (A - A^3)^- \,\}, \tag{22.26}$$

$$\{\, (I_m - A)^-(I_m + A)^- A^- \,\} \subseteq \{\, (A - A^3)^- \,\}, \quad \{\, (I_m + A)^-(I_m - A)^- A^- \,\} \subseteq \{\, (A - A^3)^- \,\}. \tag{22.27}$$

Chapter 23

Generalized inverses of sums of matrices

In this chapter we establish some rank equalities related for sums of inner inverses of matrices and then use them to deal with the following several problems:

(I) The relationship between $A^- + B^-$ and $(A + B)^-$.

(II) The relationship between $A_1^- + A_2^- + \cdots + A_k^-$ and $(A_1 + A_2 + \cdots + A_k)^-$.

(III) The relationship between $\{A^- + B^-\}$ and $\{C^-\}$.

(IV) The relationship between $\{A_1^- + A_2^- + \cdots + A_k^-\}$ and $\{C^-\}$.

We first present a formula for the dimension of the intersection of k matrices, which will be applied in the sequel.

Lemma 23.1[140]. *Let* $[A_1, A_2, \cdots A_k] \in \mathcal{F}^{m\times n}$. *Then*

$$\dim[R(A_1) \cap R(A_2) \cap \cdots \cap R(A_k)] = r(N) + r(Q) - r[N, Q], \tag{23.1}$$

where $N = \mathrm{diag}(A_1, A_2, \cdots A_k)$, $Q = [I_m, I_m, \cdots I_m]^T$. *In particular,*

$$R(A_1) \cap R(A_2) \cap \cdots \cap R(A_k) = \{0\} \Leftrightarrow R(N) \cap R(Q) = \{0\}. \tag{23.2}$$

Proof. Let $X \in \mathcal{F}^{m\times t}$ be a matrix satisfying $R(X) = \cap_{i=1}^k R(A_i)$. The this X can be written as $X = A_1X_1 = A_2X_2 = \cdots = A_kX_k$. Consider it as a system of matrix equations. It can equivalently be written as

$$\begin{bmatrix} I_m & -A_1 & & & \\ I_m & & -A_2 & & \\ \vdots & & & \ddots & \\ I_m & & & & -A_m \end{bmatrix} \begin{bmatrix} X \\ X_1 \\ \vdots \\ X_k \end{bmatrix} = 0,$$

or briefly $[Q, -N]Y = 0$. Solving for X, we obtain its general solution is

$$X = [I_m, 0](I - [Q, -N]^-[Q, -N])V,$$

where V is arbitrary. The maximal rank of X, according to (1.3) is

$$\begin{aligned} r(X) &= r([I_m, 0](I - [Q, -N]^-[Q, -N])) \\ &= r\begin{bmatrix} I_m & 0 \\ Q & -N \end{bmatrix} - r[Q, -N] \\ &= m + r(N) - r[Q, -N] = r(N) + r(Q) - r[N, Q], \end{aligned}$$

which is the dimension of $\cap_{i=1}^{k} R(A_i)$. $\square$

23.1. The relationships between $A^- + B^-$ and $(A+B)^-$

We first establish some rank equalities related to $A+B$ and $A^- + B^-$.

Theorem 23.2. *Let $A,\ B \in \mathcal{F}^{m\times n}$ be given, and let $M = A+B$. Then*

$$\max_{A^-,\, B^-} r[\, M - M(A^- + B^-)M\,] = \min\left\{ r(M), \quad r\begin{bmatrix} M & A \\ B & M \end{bmatrix} + r(M) - r(A) - r(B) \right\}, \tag{23.3}$$

and

$$\min_{A^-,\, B^-} r[\, M - M(A^- + B^-)M\,]$$
$$= r(A) + r(B) + r(M) + \ r\begin{bmatrix} M & A \\ B & M \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & B & A \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ B & A \end{bmatrix}. \tag{23.4}$$

Proof. We first show that

$$\{\, A^- + B^- \,\} = \left\{ [\, I_n,\ I_n\,]N^- \begin{bmatrix} I_m \\ I_m \end{bmatrix} \right\}, \quad \text{where} \quad N = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}. \tag{23.5}$$

In fact, the general expression of $A^- + B^-$ can be written as

$$A^- + B^- = A^{\sim} + B^{\sim} + F_A V_1 + V_2 E_A + F_B W_1 + W_2 E_B, \tag{23.6}$$

where $A^{\sim}$ and $B^{\sim}$ are two special inner inverses of A and B, V_1, V_2, W_1, and W_2 are arbitrary. The general expression of N^- is

$$\begin{aligned} N^- &= N^{\sim} + F_N S + T E_N \\ &= \begin{bmatrix} A^{\sim} & 0 \\ 0 & B^{\sim} \end{bmatrix} + \begin{bmatrix} F_A & 0 \\ 0 & F_B \end{bmatrix}\begin{bmatrix} S_1 & S_2 \\ S_3 & S_4 \end{bmatrix} + \begin{bmatrix} T_1 & T_2 \\ T_3 & T_4 \end{bmatrix}\begin{bmatrix} E_A & 0 \\ 0 & E_B \end{bmatrix} \\ &= \begin{bmatrix} A^{\sim} + F_A S_1 + T_1 E_A & F_A S_2 + T_2 E_B \\ F_B S_3 + T_3 E_A & B^{\sim} + F_B S_4 + T_4 E_B \end{bmatrix}, \end{aligned}$$

where S_1—S_4 and T_1—T_4 are arbitrary. In that case, we have the general expression

$$\begin{aligned} [\, I_n,\ I_n\,]N^- \begin{bmatrix} I_m \\ I_m \end{bmatrix} &= A^{\sim} + B^{\sim} + F_A S_1 + T_1 E_A + F_A S_2 + T_2 E_B + F_B S_3 + T_3 E_A + F_B S_4 + T_4 E_B \\ &= A^{\sim} + F_A(S_1 + S_2) + (T_1 + T_3)E_A + B^{\sim} + F_B(S_3 + S_4) + (T_2 + T_4)E_B. \end{aligned}$$

This expression is the same as (23.6). Thus (23.5) holds. This fact implies that

$$\max_{A^-,\, B^-} r[\, M - M(A^- + B^-)M\,] = \max_{N^-} r\left(M - [\, M,\ M\,]N^- \begin{bmatrix} M \\ M \end{bmatrix} \right),$$
$$\min_{A^-,\, B^-} r[\, M - M(A^- + B^-)M\,] = \min_{N^-} r\left(M - [\, M,\ M\,]N^- \begin{bmatrix} M \\ M \end{bmatrix} \right).$$

Applying (22.1) and (22.2) to the right-hand sides of the above two equalities, we obtain

$$\begin{aligned} &\max_{N^-} r\left(M - [\, M,\ M\,]N^- \begin{bmatrix} M \\ M \end{bmatrix} \right) \\ &= \min\left\{ r(M), \quad r\begin{bmatrix} A & 0 & M \\ 0 & B & M \\ M & M & M \end{bmatrix} - r(N) \right\} \\ &= \min\left\{ r(M), \quad r\begin{bmatrix} A - M & -M \\ -M & B - M \end{bmatrix} + r(M) - r(A) - r(B) \right\} \\ &= \min\left\{ r(M), \quad r\begin{bmatrix} M & A \\ B & M \end{bmatrix} + r(M) - r(A) - r(B) \right\}, \end{aligned}$$

which is exactly (23.3), and

$$\begin{aligned}
&\min_{N^-} r\left(M - [\,M,\ M\,]N^- \begin{bmatrix} M \\ M \end{bmatrix} \right) \\
&= r(N) - r\begin{bmatrix} A & 0 & M \\ 0 & B & M \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ M & M \end{bmatrix} + r\begin{bmatrix} A & 0 & M \\ 0 & B & M \\ M & M & M \end{bmatrix} \\
&= r(A) + r(B) + r(M) + r\begin{bmatrix} M & A \\ B & M \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & B & A \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ B & A \end{bmatrix},
\end{aligned}$$

which is exactly (23.4). $\square$

Two direct consequences can be derived from (23.3) and (23.4).

Theorem 23.3. *Let $A,\, B \in \mathcal{F}^{m\times n}$ be given, and let $M = A + B$. Then there exist $A^- \in \{A^-\}$ and $B^- \in \{B^-\}$ such that $A^- + B^- \in \{\,(\,A + B\,)^-\,\}$ holds if and only if*

$$r\begin{bmatrix} M & A \\ B & M \end{bmatrix} = r\begin{bmatrix} A & 0 \\ 0 & B \\ B & A \end{bmatrix} + r\begin{bmatrix} A & 0 & B \\ 0 & B & A \end{bmatrix} - r(M) - r(A) - r(B).$$

Theorem 23.4. *Let $A,\, B \in \mathcal{F}^{m\times n}$ be given, and let $M = A + B \neq 0$. The the following four statements are equivalent:*

(a) $\{\,A^- + B^-\,\} \subseteq \{\,(\,A + B\,)^-\,\}$.

(b) $r\begin{bmatrix} M & A \\ B & M \end{bmatrix} = r(A) + r(B) - r(\,A + B\,)$.

(c) $r\begin{bmatrix} A & 0 & M \\ 0 & B & M \\ M & M & M \end{bmatrix} = r\begin{bmatrix} A & 0 & B \\ 0 & B & A \\ B & A & -2M \end{bmatrix} = r\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$.

(d) $R(A) = R(B),\ R(A^T) = R(B^T)$ *and* $A + B = -\frac{1}{2}(\,AB^-A + BA^-B\,)$.

Proof. The equivalence of Parts (a) and (b) follows immediately from (22.3). The equivalence of Parts (b) and (c) follows from the rank equality

$$r\begin{bmatrix} A & 0 & M \\ 0 & B & M \\ M & M & M \end{bmatrix} = r\begin{bmatrix} M & A \\ B & M \end{bmatrix} + r(M).$$

The equivalence of Parts (c) and (d) follows from (1.5). $\square$

Theorem 23.5. *Let $A,\, B \in \mathcal{F}^{m\times n}$ be given, and let $M = A + B$. Then*

$$\min_{A^-,\, B^-} r(\,M^- - A^- - B^-\,) = r(\,M - AM^-B\,) - r\begin{bmatrix} A \\ B \end{bmatrix} - r[\,A,\ B\,] + r(A) + r(B), \tag{23.7}$$

$$\begin{aligned}
&\min_{M^-,\, A^-,\, B^-} r(\,M^- - A^- - B^-\,) \\
&= r(M) + r(A) + r(B) + r\begin{bmatrix} M & A \\ B & M \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & B & A \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ B & A \end{bmatrix},
\end{aligned} \tag{23.8}$$

$$\begin{aligned}
\max_{M^-} \min_{A^-,\, B^-} r(\,M^- - A^- - B^-\,) = \min \Bigg\{ & r(A) + r(B) - r\begin{bmatrix} A \\ B \end{bmatrix}, \qquad r(A) + r(B) - r[\,A,\ B\,], \\
& r\begin{bmatrix} M & A \\ B & M \end{bmatrix} - r\begin{bmatrix} A \\ B \end{bmatrix} - r[\,A,\ B\,] - r(M) + r(A) + r(B) \Bigg\}.
\end{aligned} \tag{23.9}$$

Proof. According to (23.5) and (22.2) we first find

$$\min_{A^-,\, B^-} r(\,M^- - A^- - B^-\,)$$

$$
\begin{aligned}
&= \min_{N^-}\left(M^- - [\, I_n,\ I_n\,]N^- \begin{bmatrix} I_m \\ I_m \end{bmatrix}\right) \\
&= r(N) - r\begin{bmatrix} A & 0 & I_m \\ 0 & B & I_m \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ I_n & I_n \end{bmatrix} + r\begin{bmatrix} A & 0 & I_m \\ 0 & B & I_m \\ I_n & I_n & M^- \end{bmatrix} \\
&= r(N) - r\begin{bmatrix} 0 & 0 & I_m \\ -A & B & 0 \end{bmatrix} - r\begin{bmatrix} 0 & -A \\ 0 & B \\ I_n & 0 \end{bmatrix} + r\begin{bmatrix} 0 & -M + AM^-B & 0 \\ 0 & 0 & I_m \\ I_n & 0 & 0 \end{bmatrix} \\
&= r(\, M - AM^-B\,) - r\begin{bmatrix} A \\ B \end{bmatrix} - r[\, A,\ B\,] + r(A) + r(B),
\end{aligned}
$$

which is eaxactly (23.7). Next applying (21.4) and (21.5) to $M - AM^-B$, we obtain

$$
\begin{aligned}
&\min_{M^-} r(\, M - AM^-B\,) \\
&= r(M) + r[\, A,\ M\,] + r\begin{bmatrix} B \\ M \end{bmatrix} + r\begin{bmatrix} M & B \\ A & M \end{bmatrix} - r\begin{bmatrix} M & 0 & B \\ 0 & A & M \end{bmatrix} - r\begin{bmatrix} M & 0 \\ 0 & B \\ A & M \end{bmatrix} \\
&= r(M) + r[\, A,\ B\,] + r\begin{bmatrix} A \\ B \end{bmatrix} + r\begin{bmatrix} M & A \\ B & M \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & B & A \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ B & A \end{bmatrix},
\end{aligned}
$$

$$
\max_{M^-} r(\, M - AM^-B\,) = \min\left\{ r[\, A,\ B\,], \quad r\begin{bmatrix} A \\ B \end{bmatrix}, \quad r\begin{bmatrix} M & B \\ A & M \end{bmatrix} - r(M) \right\}.
$$

Putting them in (23.7) resprctively yields (23.8) and (23.9). $\square$

Two direct consequences of Theorem 23.5 are given below.

Theorem 23.6. *Let A, $B \in \mathcal{F}^{m\times n}$ be given. Then for a given $(\, A + B\,)^-$, there exist $A^- \in \{A^-\}$ and $B^- \in \{B^-\}$ such that $A^- + B^- = (\, A + B\,)^-$ if and only if*

$$
r[\, A + B - A(\, A + B\,)^-B\,] = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A,\ B\,] - r(A) - r(B).
$$

Theorem 23.7. *Let A, $B \in \mathcal{F}^{m\times n}$ be given. Then $\{(\, A + B\,)^-\} \subseteq \{\, A^- + B^-\,\}$ holds if and only if*

$$
R(A) \cap R(B) = \{0\}, \quad \text{or} \quad R(A^T) \cap R(B^T) = \{0\},
$$

or

$$
r\begin{bmatrix} A + B & A \\ B & A + B \end{bmatrix} = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A,\ B\,] + r(\, A + B\,) - r(A) - r(B).
$$

Combining Theorems 23.4 and 23.7, one can easily establish a necessary and sufficient condition for $\{(\, A + B\,)^-\} = \{\, A^- + B^-\,\}$ to hold. we shall, however, to present it in Section 23.3 as a special case of a general result.

23.2. The relationships between $A_1^- + A_2^- + \cdots + A_k^-$ and $(A_1 + A_2 + \cdots + A_k\,)^-$

The results in the preceding section can directly be extended to sums of k matrices. We present them below without detailed proofs.

Theorem 23.8. *Let A_1, A_2, $\cdots$, $A_k \in \mathcal{F}^{m\times n}$ be given, and let $M = A_1 + A_2 + \cdots + A_k$. Then*

$$
\max_{A_1^-, \cdots, A_k^-} r[\, M - M(\, A_1^- + \cdots + A_k^-\,)M\,] = \min\{\, r(M), \quad r(\, N - QMP\,) + r(M) - r(N)\,\}, \tag{23.10}
$$

and

$$
\min_{A_1^-, \cdots, A_k^-} r[\, M - M(\, A_1^- + \cdots + A_k^-\,)M\,] = r(M) + r(N) + r(\, N - QMP\,) - r[\, N,\ QM\,] - r\begin{bmatrix} N \\ MP \end{bmatrix}, \tag{23.11}
$$

where $N = \mathrm{diag}(\, A_1,\, A_2,\, \cdots,\, A_k\,)$, $P = [\, I_n,\, I_n,\, \cdots\, I_n\,]$, $Q = [\, I_m,\, I_m,\, \cdots\, I_m\,]^T$.

Proof. It is easy to verify that

$$\{\, A_1^- + \cdots + A_k^- \,\} = \{\, PN^-Q \,\}. \tag{23.12}$$

In that case, it follows by (22.1) that

$$\begin{aligned}
&\max_{A_1^-, \cdots, A_k^-} r[\, M - M(\, A_1^- + \cdots + A_k^- \,)M \,] \\
&= \max_{N^-} r(\, M - MPN^-QM \,) \\
&= \min\left\{ r(M), \quad r\begin{bmatrix} N & QM \\ MP & M \end{bmatrix} - r(N) \right\} \\
&= \min\{\, r(M), \quad r(\, N - QMP \,) + r(M) - r(N) \,\},
\end{aligned}$$

which is exactly (3.1). Applying (22.2), we also obtain

$$\begin{aligned}
&\min_{A_1^-, \cdots, A_k^-} r[\, M - M(\, A_1^- + \cdots + A_k^- \,)M \,] \\
&= \min_{N^-} r(\, M - MPN^-QM \,) \\
&= r(N) - r[\, N,\ QM \,] - r\begin{bmatrix} N \\ MP \end{bmatrix} + r\begin{bmatrix} N & QM \\ MP & M \end{bmatrix} \\
&= r(N) - r[\, N,\ QM \,] - r\begin{bmatrix} N \\ MP \end{bmatrix} + r(\, N - QMP \,) + r(M),
\end{aligned}$$

which is exactly (23.11). □

Two direct consequences of (23.10) and (23.11) are listed below.

Theorem 23.9. *Let* $A_1,\, A_2,\, \cdots,\, A_k \in \mathcal{F}^{m\times n}$ *be given, and denote* $M = A_1 + A_2 + \cdots + A_k$. *Then there exist* $A_i^- \in \{A_i^-\}$, $i = 1,\, 2,\, \cdots,\, k$ *such that* $A_1^- + A_2^- + \cdots + A_k^- \in \{M^-\}$ *if and only if*

$$r(\, N - QMP \,) = r\begin{bmatrix} N \\ MP \end{bmatrix} + r[\, N,\ QM \,] - r(M) - r(N),$$

where $N,\ P$ *and* Q *are defined in Theorem 23.8.*

Theorem 23.10. *Let* $A_1,\, A_2,\, \cdots,\, A_k \in \mathcal{F}^{m\times n}$ *be given, and let* $M = A_1 + A_2 + \cdots + A_k$. *The the following four statements are equivalent*:

(a) $\{\, A_1^- + A_2^- \cdots + A_k^- \,\} \subseteq \{\, (\, A_1 + A_2 + \cdots + A_k \,)^- \,\}$.

(b) $r(\, N - QMP \,) = r(N) - r(QMP)$.

(c) $r\begin{bmatrix} N & QM \\ MP & M \end{bmatrix} = r(N)$.

(d) $R(M) = R(A_i)$, $R(M^T) = R(A_i^T)$, $i = 1,\, 2,\, \cdots,\ k$, *and* $M = MPN^-QM$, *where* $N,\ P$ *and* Q *are defined in Theorem 23.8.*

Proof. The equivalence of Parts (a) and (b) follows immediately from (23.10). The equivalence of Parts (b) and (c) is evident. The equivalence of Parts (c) and (d) follows from (1.5). □

Theorem 23.11. *Let* $A_1,\, A_2,\, \cdots,\, A_k \in \mathcal{F}^{m\times n}$ *be given, and denote* $M = A_1 + A_2 + \cdots + A_k$. *Then*

$$\min_{A_1^-, \cdots, A_k^-} r(\, M^- - A_1^- - \cdots - A_k^- \,) = r(N) - r[\, N,\ Q \,] - r\begin{bmatrix} N \\ P \end{bmatrix} + r\begin{bmatrix} N & Q \\ P & M^- \end{bmatrix}, \tag{23.13}$$

$$\min_{M^-, A_1^-, \cdots, A_k^-} r(\, M^- - A_1^- - \cdots - A_k^- \,) = r(M) + r(N) + r(\, N - QMP \,) - r[\, N,\ QM \,] - r\begin{bmatrix} N \\ MP \end{bmatrix}, \tag{23.14}$$

$$\begin{aligned}
\max_{M^-} \min_{A_1^-, \cdots, A_k^-} r(\, M^- - A_1^- - \cdots - A_k^- \,) = \min\Bigg\{ & r(N) + m - r[\, N,\ Q \,], \quad r(N) + n - r\begin{bmatrix} N \\ P \end{bmatrix}, \\
& m + n + r(\, N - QMP \,) + r(N) - r(M) - r[\, N,\ Q \,] - r\begin{bmatrix} N \\ P \end{bmatrix} \Bigg\},
\end{aligned} \tag{23.15}$$

where $N,\ P$ and Q are defined in Theorem 23.8.

Proof. According to (23.12) and (22.2) we first find that

$$\begin{aligned}&\min_{A_1^-,\,\cdots,\,A_k^-} r(\,M^- - A_1^- - \cdots - A_k^-\,)\\ &= \min_{N^-}(\,M^- - PN^-Q\,) = r(N) - r[\,N,\ Q\,] - r\begin{bmatrix} N \\ P\end{bmatrix} + r\begin{bmatrix} N & Q \\ P & M^-\end{bmatrix},\end{aligned}$$

which is eaxactly (23.13). Applying (21.5) to the 2×2 block matrix in the above equality, we further obtain

$$\begin{aligned}&\min_{M^-} r\begin{bmatrix} N & Q \\ P & M^-\end{bmatrix}\\
=\ & \min_{M^-} r\left(\begin{bmatrix} N & Q \\ P & 0\end{bmatrix} + \begin{bmatrix} 0 \\ I_n\end{bmatrix} M^-[\,0,\ I_m\,]\right)\\
=\ & r(M) + r\begin{bmatrix} N & Q & 0 \\ P & 0 & I_n\end{bmatrix} + r\begin{bmatrix} N & Q \\ P & 0 \\ 0 & I_m\end{bmatrix} + r\begin{bmatrix} -M & 0 & I_m \\ 0 & N & Q \\ I_n & P & 0\end{bmatrix}\\
&-\ r\begin{bmatrix} -M & 0 & 0 & I_m \\ 0 & 0 & N & Q \\ 0 & I_n & P & 0\end{bmatrix} - r\begin{bmatrix} -M & 0 & 0 \\ 0 & 0 & I_m \\ 0 & N & Q \\ I_n & P & 0\end{bmatrix}\\
=\ & r(M) + r[\,N,\ Q\,] + r\begin{bmatrix} N \\ P\end{bmatrix} + r\begin{bmatrix} 0 & 0 & I_m \\ 0 & N - QMP & 0 \\ I_n & 0 & 0\end{bmatrix} - r\begin{bmatrix} M & 0 & I_m \\ 0 & N & Q\end{bmatrix} - r\begin{bmatrix} M & 0 \\ 0 & N \\ I_n & P\end{bmatrix}\\
=\ & r(M) + r[\,N,\ Q\,] + r\begin{bmatrix} N \\ P\end{bmatrix} + r(\,N - QMP\,) - r[\,N,\ QM\,] - r\begin{bmatrix} N \\ MP\end{bmatrix}.\end{aligned}$$

Putting it in (23.13) yields (23.14). Next applying (21.5), we obtain the following

$$\begin{aligned}\max_{M^-} r\begin{bmatrix} N & Q \\ P & M^-\end{bmatrix} &= \max_{M^-} r\left(\begin{bmatrix} N & Q \\ P & 0\end{bmatrix} + \begin{bmatrix} 0 \\ I_n\end{bmatrix} M^-[\,0,\ I_m\,]\right)\\
&= \min\left\{ r\begin{bmatrix} N & Q & 0 \\ P & 0 & I_n\end{bmatrix},\quad r\begin{bmatrix} N & Q \\ P & 0 \\ 0 & I_n\end{bmatrix},\quad r\begin{bmatrix} -M & 0 & I_m \\ 0 & N & Q \\ I_n & P & 0\end{bmatrix} - r(M)\right\}\\
&= \min\left\{ n + r[\,N,\ Q\,],\quad m + r\begin{bmatrix} N \\ P\end{bmatrix},\quad m + n + r(\,N - QMP\,) - r(M)\right\}.\end{aligned}$$

Putting it in (23.13) yields (23.15). □

Two direct consequences of Theorem 23.11 are given below.

Theorem 23.12. *Let $A_1,\, A_2,\, \cdots,\, A_k \in \mathcal{F}^{m\times n}$ be given, and let $M = A_1 + A_2 + \cdots + A_k$. Then for a given M^-, there exist $A_i^- \in \{A_i^-\}$, $i = 1,\, 2,\, \cdots,\, k$ such that $A_1^- + A_2^- + \cdots + A_k^- = M^-$ if and only if M^- satisfies*

$$r\begin{bmatrix} N & Q \\ P & M^-\end{bmatrix} = r\begin{bmatrix} N \\ P\end{bmatrix} + r[\,N,\ Q\,] - r(N),$$

where $N,\ P$ and Q are defined in Theorem 23.8.

Theorem 23.13. *Let $A_1,\, A_2,\, \cdots,\, A_k \in \mathcal{F}^{m\times n}$ be given, and let $M = A_1 + A_2 + \cdots + A_k$. Then the set inclusion*

$$\{\,(\,A_1 + A_2 + \cdots + A_k\,)^-\,\} \subseteq \{\,A_1^- + A_2^- + \cdots + A_k^-\,\} \tag{23.16}$$

holds if and only if

$$R(A_1) \cap R(A_2) \cap \cdots \cap R(A_k) = \{0\}, \tag{23.17}$$

or

$$R(A_1^T) \cap R(A_2^T) \cap \cdots \cap R(A_k^T) = \{0\}, \tag{23.18}$$

or

$$r(\, N - QMP\,) = r\begin{bmatrix} N \\ P \end{bmatrix} + r[\, N,\ Q\,] - r(N) + r(M) - m - n, \tag{23.19}$$

where $N,\ P$ and Q are defined in Theorem 23.8.

Proof. It is easy to see that the set inclusion in (23.16) hold if and only if

$$\max_{M^-} \min_{A_1^-, \cdots, A_k^-} r(\, M^- - A_1^- - \cdots - A_k^-\,) = 0.$$

In light of (23.15), the above equality is equivalent to

$$r[\, N,\ Q\,] = r(N) + r(Q), \quad \text{or} \quad r\begin{bmatrix} N \\ P \end{bmatrix} = r(N) + r(P), \tag{23.20}$$

or (23.18) holds. The two rank equalities in (23.20) are equivalent to (23.17) and (23.18) according to (23.1). □

23.3. The relationships between $\{\, A^- + B^-\,\}$ and $\{\, C^-\,\}$, and parallel sum of two matrices

It is well known (see [107], [118]) that the parallel sum of two matrices A and B of the same size is defined to be $C := A(\, A + B\,)^- B$, whenever this product is invariant with respect to the choice of $(\, A + B\,)^-$. One of the well-known nice properties on parallel sum of two matrices is $\{\, A^- + B^-\,\} = \{C^-\}$. This set equality motivates us to consider the relationship between the two sets $\{\, A^- + B^-\,\}$ and $\{C^-\}$ in general cases, where $A,\ B$, and C are any three given matrices of the same size. Just as what we do in Section 23.1, we first establish several basic rank equalities related to generalized inverses of $A,\ B$, and C, and then deduce from them various relationships between $\{\, A^- + B^-\,\}$ and $\{\, C^-\}$.

Theorem 23.14. *Let $A,\, B,\, C \in \mathcal{F}^{m\times n}$ be given. Then*

$$\max_{A^-,\, B^-} r[\, C - C(\, A^- + B^-\,)C\,] = \min\left\{ r(C), \quad r\begin{bmatrix} A - C & C \\ C & B - C \end{bmatrix} + r(C) - r(A) - r(B) \right\}, \tag{23.21}$$

$$\min_{A^-,\, B^-} r[\, C - C(\, A^- + B^-\,)C\,] = r(A) + r(B) + r(C) + r\begin{bmatrix} A - C & C \\ C & B - C \end{bmatrix} - r\begin{bmatrix} A & 0 & C \\ 0 & B & C \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & C \end{bmatrix}. \tag{23.22}$$

Proof. By (23.5) and (22.1), we easily find

$$\begin{aligned} \max_{A^-,\, B^-} r[\, C - C(\, A^- + B^-\,)C\,] &= \max_{N^-} r\left(C - [\, C,\ C\,]N^- \begin{bmatrix} C \\ C \end{bmatrix} \right) \\ &= \min\left\{ r(C), \quad r\begin{bmatrix} A & 0 & C \\ 0 & B & C \\ C & C & C \end{bmatrix} - r(N) \right\} \\ &= \min\left\{ r(C), \quad r\begin{bmatrix} A - C & -C \\ -C & B - C \end{bmatrix} + r(C) - r(A) - r(B) \right\}, \end{aligned}$$

which is (23.21). Next applying (23.5) and (22.2), we obtain

$$\min_{A^-,\, B^-} r[\, C - C(\, A^- + B^-\,)C\,]$$

$$= \min_{N^-} r\left(C - [C,\ C]N^- \begin{bmatrix} C \\ C \end{bmatrix} \right)$$

$$= r(N) - r\begin{bmatrix} A & 0 & C \\ 0 & B & C \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & C \end{bmatrix} + r\begin{bmatrix} A & 0 & C \\ 0 & B & C \\ C & C & C \end{bmatrix}$$

$$= r(A) + r(B) + r(C) + r\begin{bmatrix} A - C & C \\ C & B - C \end{bmatrix} - r\begin{bmatrix} A & 0 & C \\ 0 & B & C \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & C \end{bmatrix},$$

which is exactly (23.22). □

Two consequences can directly be derived from (23.21) and (23.22).

Theorem 23.15. *Let $A,\ B,\ C \in \mathcal{F}^{m\times n}$ be given. Then there exist $A^- \in \{A^-\}$ and $B^- \in \{B^-\}$ such that $A^- + B^- \in \{\, C^- \,\}$ holds if and only if*

$$r\begin{bmatrix} A - C & C \\ C & B - C \end{bmatrix} = r\begin{bmatrix} C & B \\ A & A + B \end{bmatrix} = r\begin{bmatrix} A & 0 \\ 0 & B \\ B & A \end{bmatrix} + r\begin{bmatrix} A & 0 & B \\ 0 & B & A \end{bmatrix} - r(N) - r(A) - r(B). \quad (23.23)$$

Theorem 23.16. *Let $A,\ B,\ C \in \mathcal{F}^{m\times n}$ be given with $C \neq 0$. Then the following four statements are equivalent*:

(a) $\{\, A^- + B^- \,\} \subseteq \{\, C^- \}$.

(b) $r\begin{bmatrix} A - C & C \\ C & B - C \end{bmatrix} = r\begin{bmatrix} C & B \\ A & A + B \end{bmatrix} = r(A) + r(B) - r(C).$

(c) $r\begin{bmatrix} A & 0 & C \\ 0 & B & C \\ C & C & C \end{bmatrix} = r\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$

(d) $R(C) \subseteq R(A),\ R(C) \subseteq R(B),\ R(C^T) \subseteq R(A^T),\ R(C^T) \subseteq r(B^T)$ *and* $C = CA^-C + CB^-C.$

Proof. The equivalence of Parts (a) and (b) follows from (23.21). The equivalence of Parts (b) and (c) is evident. The equivalence of Parts (c) and (d) follows from (1.5). □

Theorem 23.17. *Let $A,\ B,\ C \in \mathcal{F}^{m\times n}$ be given. Then*

$$\min_{A^-,\,B^-} r(\, C^- - A^- - B^- \,) = r(\, A + B - AC^-B \,) - r\begin{bmatrix} A \\ B \end{bmatrix} - r[\, A,\ B\,] + r(A) + r(B), \quad (23.24)$$

$$\min_{C^-,\,A^-,\,B^-} r(\, C^- - A^- - B^- \,) = r(C) + r(A) + r(B) + r\begin{bmatrix} C & B \\ A & A + B \end{bmatrix} - r\begin{bmatrix} A & 0 & C \\ 0 & B & C \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & C \end{bmatrix}, \quad (23.25)$$

$$\max_{C^-}\ \min_{A^-,\,B^-} r(\, C^- - A^- - B^- \,) = \min\left\{ r(A) + r(B) - r\begin{bmatrix} A \\ B \end{bmatrix},\quad r(A) + r(B) - r[\, A,\ B\,], \right.$$
$$\left. r\begin{bmatrix} C & B \\ A & A + B \end{bmatrix} - r\begin{bmatrix} A \\ B \end{bmatrix} - r[\, A,\ B\,] - r(C) + r(A) + r(B) \right\}. \quad (23.26)$$

The proof of this theorem is much similar to that of Theorem 23.5 and is, therefore, omitted. Two direct consequences of Theorem 23.17 are given below.

Theorem 23.18. *Let $A,\ B,\ C \in \mathcal{F}^{m\times n}$ be given. Then for a given C^-, there exist $A^- \in \{A^-\}$ and $B^- \in \{B^-\}$ such that $A^- + B^- = C^-$ if and only if $A,\ B$ and C^- satisfies*

$$r[\, A + B - AC^-B \,] = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A,\ B\,] - r(A) - r(B). \quad (23.27)$$

Theorem 23.19. *Let $A,\ B,\ C \in \mathcal{F}^{m\times n}$ be given. Then $\{C^-\} \subseteq \{\, A^- + B^- \,\}$ holds if and only if*

$$R(A)\cap R(B) = \{0\}, \quad or \quad R(A^T)\cap R(B^T) = \{0\}, \tag{23.28}$$

or

$$r\begin{bmatrix} C & B \\ A & A+B \end{bmatrix} = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A,\ B\,] + r(C) - r(A) - r(B). \tag{23.29}$$

The two conditions in (23.28) have no relation with the matrix C, which implies that under (23.28), $\{C^-\} \subseteq \{\, A^- + B^- \,\}$ holds for any choice of C. Now setting $C = 0$, then $\{\, 0^- \} = \mathcal{F}^{n\times m}$. Thus under (23.28), there is

$$\{\, A^- + B^- \,\} = \{\, 0^- \} = \mathcal{F}^{n\times m}. \tag{23.30}$$

Conversely, if (23.30) holds, then it is easy to see from Theorem 23.19 that (23.28) holds. Thus (23.28) is a necessary and sufficient condition for (23.30) to hold.

Now combining Theorems 23.16 and 23.19, we obtain the following result.

Theorem 23.20. *Let $A,\, B,\, C \in \mathcal{F}^{m\times n}$ be given and suppose that*

$$R(A)\cap R(B) \neq \{0\}, \quad and \quad R(A^T)\cap R(B^T) \neq \{0\}. \tag{23.31}$$

Then the equality

$$\{\ A^- + B^- \,\} = \{C^-\} \tag{23.32}$$

holds if and only if

$$R(B) \subseteq R(\, A + B\,), \quad R(A^T) \subseteq R(\, A^T + B^T\,) \ \ and \ \ C = A(\, A + B\,)^- B, \tag{23.33}$$

that is, A and B are parallel summable and C is the parallel sum of $A$and B. In that case, the rank of C satisfies the following rank equality

$$r(C) = r(A) + r(B) - r(\, A + B\,). \tag{23.34}$$

Proof. Assume first that (23.32) holds. By Theorem 23.19, we know that $A,\ B$ and C satisfy (23.29). On the other hand, (23.32) implies that

$$\min_{A^-,\,B^-} r(\, A^- + B^-\,) = \min_{C^-} r(C^-). \tag{23.35}$$

It is well-known that the minimal rank of C^- is $r(C)$. On the other hand, it follows from (23.24) that

$$\min_{A^-,\,B^-} r(\, A^- + B^-\,) = r(\, A + B\,) + r(A) + r(B) - r[\, A,\ B\,] - r\begin{bmatrix} A \\ B \end{bmatrix}.$$

Thus (23.35) is equivalent to

$$r(C) = r(\, A + B\,) + r(A) + r(B) - r[\, A,\ B\,] - r\begin{bmatrix} A \\ B \end{bmatrix}. \tag{23.36}$$

Putting it (23.29) yields

$$r\begin{bmatrix} C & B \\ A & A+B \end{bmatrix} = r(\, A + B\,), \tag{23.37}$$

which, by (1.5), is equivalent to (23.33), meanwhile

$$\begin{aligned} r(C) &= \min_{(A+B)^-} r[\, A(\, A + B\,)^- B\,] \\ &= r(\, A + B\,) - r[\, A + B,\ B\,] - r\begin{bmatrix} A+B \\ B \end{bmatrix} + r\begin{bmatrix} A+B & B \\ A & 0 \end{bmatrix} \\ &= r(A) + r(B) - r(\, A + B\,). \end{aligned}$$

Conversely, if (23.33) holds, then (23.34) and (23.37) also hold. Combining both of them shows that the two rank equalities in Theorem 23.16(b) and (23.29) are satisfied. Therefore (23.32) holds. □

The equivalence of (23.32) and (23.33) was previously proved by Mitra and Odell in [107]. But the assumption (23.31) is neglected there. As shown in (23.30), if $R(A) \cap R(B) = \{0\}$ and $R(A^T) \cap R(B^T) \neq \{0\}$, then the equality $\{ A^- + B^- \} = \{ 0^- \}$. This, however, does not imply that A and B are parallel summable and $A(A + B)^- B = 0$, in general. An example is

$$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

both of which satisfy $\{ A^- + B^- \} = \{ 0^- \}$, since $R(A) \cap R(B) = \{0\}$. If we let

$$(A + B)^- = \begin{bmatrix} 1 \\ 1 \end{bmatrix}^- = \frac{1}{2}[1,\ 1],$$

then

$$A(A + B)^- B = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} [1,\ 1] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq 0.$$

If we let

$$(A + B)^- = \begin{bmatrix} 1 \\ 1 \end{bmatrix}^- = [1,\ 0],$$

then

$$A(A + B)^- B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} [1,\ 0] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Therefore A and B are not parallel summable.

Some interesting consequences can be derived from the above results. For example, let $B = I_m - A$ and $C = I_m$ in (23.24), we then get

$$\min_{A^-,\, (I_m - A)^-} r[\, I_m - A^- - (I_m - A)^- \,] = r[\, (A - A^2) - (A - A^2)^2 \,].$$

Thus there are A^- and $(I_m - A)^-$ such that $A^- + (I_m - A)^- = I_m$ if and only if $A - A^2$ is idempotent.

Replace $B = I_m - A$ and $C = A - A^2$ in Theorem 23.20. Then it is easy to verify that these A, B and C satisfy the condition (23.33). Thus the set equality

$$\{ (A - A^2)^- \} = \{ A^- + (I_m - A)^- \}$$

holds for any A. In other words, the matrices A and $I_m - A$ are always parallel summable, and $A - A^2$ is their parallel sum. Recall (22.23), we then get the following

$$\{ A^-(I_m - A)^- \} \subseteq \{ A^- + (I_m - A)^- \} \quad \text{and} \quad \{ (I_m - A)^- A^- \} \subseteq \{ A^- + (I_m - A)^- \}.$$

When one of A and $I_m - A$ is nonsingular, say, A, there is

$$\{ A^{-1}(I_m - A)^- \} \subseteq \{ A^{-1} + (I_m - A)^- \} \quad \text{and} \quad \{ (I_m - A)^- A^{-1} \} \subseteq \{ A^{-1} + (I_m - A)^- \}.$$

When both A and $I_m - A$ nonsingular, the above becomes a trivial result $A^{-1}(I_m - A)^{-1} = A^{-1} + (I_m - A)^{-1}$.

Replace A, B and C in Theorem 23.20 by $I_m + A$, $I_m - A$ and $(I_m - A^2)/2$, respectively. Then it is easy to verify that they satisfy the condition (23.33). Thus the set equality

$$\{ (I_m - A^2)^- \} = \left\{ \frac{1}{2}(I_m + A)^- + \frac{1}{2}(I_m - A)^- \right\}$$

holds for any A. In other words, the matrices $I_m + A$ and $I_m - A$ are always parallel summable, and the matrix $(I_m - A^2)/2$ is their parallel sum. Recall (22.24), then we also get the following

$$\{ (I_m + A)^-(I_m - A)^- \} \subseteq \left\{ \frac{1}{2}(I_m + A)^- + \frac{1}{2}(I_m - A)^- \right\},$$

$$\{\,(I_m - A)^-(I_m + A)^-\,\} \subseteq \left\{\,\frac{1}{2}(I_m + A)^- + \frac{1}{2}(I_m - A)^-\,\right\}.$$

When one of $I_m + A$ and $I_m - A$ is nonsingular, say, $I_m + A$, there is

$$\{\,(I_m + A)^{-1}(I_m - A)^-\,\} \subseteq \left\{\,\frac{1}{2}(I_m + A)^{-1} + \frac{1}{2}(I_m - A)^-\,\right\},$$

$$\{\,(I_m - A)^-(I_m + A)^{-1}\,\} \subseteq \left\{\,\frac{1}{2}(I_m + A)^{-1} + \frac{1}{2}(I_m - A)^-\,\right\}.$$

When both $I_m + A$ and $I_m - A$ are nonsingular, the above becomes a trivial result $2(I_m + A)^{-1}(I_m - A)^{-1} = (I_m + A)^{-1} + (I_m - A)^{-1}$.

In general, suppose that $\lambda_1 \neq \lambda_2$ are two scalars, and replacing A, B and C in Theorem 23.20 by $\lambda_1 I_m - A$, $\lambda_2 I_m - A$ and $(\lambda_1 I_m - A)(\lambda_2 I_m - A)/(\lambda_1 - \lambda_2)$, respectively. Then it is easy to verify that they satisfy the condition (23.33). Thus the set equality

$$\{\,[\,(\lambda_1 I_m - A)(\lambda_2 I_m - A)\,]^-\,\} = \left\{\,\frac{1}{\lambda_1 - \lambda_2}(\lambda_1 I_m - A)^- + \frac{1}{\lambda_2 - \lambda_1}(\lambda_2 I_m - A)^-\,\right\}$$

holds for any A. In other words, the matrices $\lambda_1 I_m - A$ and $\lambda_2 I_m - A$ are parallel summable, and the matrix $(\lambda_1 I_m - A)(\lambda_2 I_m - A)/(\lambda_1 - \lambda_2)$ is their parallel sum. Recall (22.22), we then get the following

$$\{\,(\lambda_1 I_m - A)^-(\lambda_2 I_m - A)^-\,\} \subseteq \left\{\,\frac{1}{\lambda_1 - \lambda_2}(\lambda_1 - I_m A)^- + \frac{1}{\lambda_2 - \lambda_1}(\lambda_2 I_m - A)^-\,\right\}.$$

This result motivates us to guess that for $\lambda_1, \cdots, \lambda_k$ with $\lambda_i \neq \lambda_j$ for all $i \neq j$,there is

$$\{(\lambda_1 I_m - A)^- \cdots (\lambda_k I_m - A)^-\} \subseteq \left\{\frac{1}{p_1}(\lambda_1 - I_m A)^- + \cdots + \frac{1}{p_k}(\lambda_k I_m - A)^-\right\},$$

where

$$p_i = (\lambda_1 - \lambda_i)\cdots(\lambda_{i-1} - \lambda_i)(\lambda_{i+1} - \lambda_i)\cdots(\lambda_k - \lambda_i), \quad i = 1, 2, \cdots, k.$$

We leave it as an open problem to the reader.

23.4. The relationships between $\{\,A_1^- + A_2^- + \cdots + A_k^-\,\}$ and $\{C^-\}$, and parallel sum of k matrices

The results in Section 23.3 can easily be generalized to sums of k matrices, which can help to extend the concept of the parallel sum of two matrices to k matrices, and establish a set of results on parallel sums of k matrices.

Theorem 23.21. *Let $A_1, A_2, \cdots, A_k, C \in \mathcal{F}^{m \times n}$ be given. Then*

$$\max_{A_1^-, \cdots, A_k^-} r[\,C - C(\,A_1^- + \cdots + A_k^-\,)C\,] = \min\{\,r(C), \quad r(\,N - QCP\,) + r(C) - r(N)\,\}, \tag{23.38}$$

$$\min_{A_1^-, \cdots, A_k^-} r[\,C - C(\,A_1^- + \cdots + A_k^-\,)C\,] = r(C) + r(N) + r(\,N - QCP\,) - r[\,N, \ QC\,] - r\begin{bmatrix} N \\ CP \end{bmatrix}, \tag{23.39}$$

where $N = \mathrm{diag}(\,A_1, A_2, \cdots, A_k\,)$, $P = [\,I_n, I_n, \cdots, I_n\,]$, and $Q = [\,I_m, I_m, \cdots, I_m\,]^T$.

Proof. According to (23.12), (22.1) and (22.2), we easily find that

$$\begin{aligned} \max_{A_1^-, \cdots, A_k^-} r[\,C - C(\,A_1^- + \cdots + A_k^-\,)C\,] &= \max_{N^-} r(\,C - CPN^-QC\,) \\ &= \min\left\{ r(C), \quad r\begin{bmatrix} N & QC \\ CP & C \end{bmatrix} - r(N) \right\} \\ &= \min\{\,r(C), \quad r(\,N - QCP\,) + r(C) - r(N)\,\}, \end{aligned}$$

$$\begin{aligned}\min_{A_1^-, \cdots, A_k^-} r[\, C - C(\, A_1^- + \cdots + A_k^- \,)C \,] &= \min_{N^-} r(\, C - CPN^-QC \,) \\ &= r(N) - r[\, N, \ QC \,] - r\begin{bmatrix} N \\ CP \end{bmatrix} + r\begin{bmatrix} N & QC \\ CP & C \end{bmatrix} \\ &= r(N) - r[\, N, \ QC \,] - r\begin{bmatrix} N \\ CP \end{bmatrix} + r(\, N - QCP \,) + r(C),\end{aligned}$$

establishing (23.38) and (23.39). □

Two consequences can directly be derived from (23.38) and (23.39).

Theorem 23.22. *Let $A_1, \, A_2, \, \cdots, \, A_k, \, C \in \mathcal{F}^{m \times n}$ be given. Then there exist $A_i^- \in \{A_i^-\}$, $i = 1, \, 2, \, \cdots, \, k$ such that $A_1^- + A_2^- + \cdots + A_k^- \in \{C^-\}$, if and only if*

$$r(\, N - QCP \,) = r\begin{bmatrix} N \\ CP \end{bmatrix} + r[\, N, \ QC \,] - r(N) - r(C),$$

where $N, \ P$ and Q are defined in Theorem 23.21.

Theorem 23.23. *Let $A_1, \, A_2, \, \cdots, \, A_k, \, C \in \mathcal{F}^{m \times n}$ be given. The the following four statements are equivalent*:

(a) $\{\, A_1^- + A_2^- \cdots + A_k^- \,\} \subseteq \{C^-\}$.

(b) $r(\, N - QCP \,) = r(N) - r(QCP)$.

(c) $r\begin{bmatrix} N & QC \\ CP & C \end{bmatrix} = r(N)$.

(d) $R(C) = r(A_i), \ R(C^T) = r(A_i^T), \ i = 1 — k,$ *and* $C = CPN^-QC$, *where $N, \ P$ and Q are defined in Theorem 23.21.*

Theorem 23.24. *Let $A_1, \, A_2, \, \cdots, \, A_k, \, C \in \mathcal{F}^{m \times n}$ be given. Then*

$$\min_{A_1^-, \cdots, A_k^-} r(\, C^- - A_1^- - \cdots - A_k^- \,) = r(N) - r[\, N, \ Q \,] - r\begin{bmatrix} N \\ P \end{bmatrix} + r\begin{bmatrix} N & Q \\ P & C^- \end{bmatrix},$$

$$\min_{C^-, A_1^-, \cdots, A_k^-} r(\, C^- - A_1^- - \cdots - A_k^- \,) = r(C) + r(N) + r(\, N - QCP \,) - r[\, N, \ QC \,] - r\begin{bmatrix} N \\ CP \end{bmatrix},$$

$$\begin{aligned}\max_{C^-} \min_{A_1^-, \cdots, A_k^-} r(\, C^- - A_1^- - \cdots - A_k^- \,) = \min \Bigg\{ & r(N) + m - r[\, N, \ Q \,], \quad r(N) + n - r\begin{bmatrix} N \\ P \end{bmatrix}, \\ & m + n + r(\, N - QCP \,) + r(N) - r(C) - r[\, N, \ Q \,] - r\begin{bmatrix} N \\ P \end{bmatrix} \Bigg\},\end{aligned}$$

where $N, \ P$ and Q are defined in Theorem 23.21.

The proof of this theorem is much like that of Theorem 23.17 and is, therefore, omitted. Two direct consequences of Theorem 23.24 are given below.

Theorem 23.25. *Let $A_1, \, A_2, \, \cdots, \, A_k, \, C \in \mathcal{F}^{m \times n}$ be given. Then for a given C^-, there exist $A_i^- \in \{A_i^-\}$, $i = 1, \, 2, \, \cdots, \, k$ such that*

$$A_1^- + A_2^- + \cdots + A_k^- = C^-,$$

if and only if C^- satisfies

$$r\begin{bmatrix} N & Q \\ P & C^- \end{bmatrix} = r\begin{bmatrix} N \\ P \end{bmatrix} + r[\, N, \ Q \,] - r(N),$$

where $N, \ P$ and Q are defined in Theorem 23.21.

Theorem 23.26. *Let $A_1,\ A_2,\ \cdots,\ A_k,\ C \in \mathcal{F}^{m\times n}$ be given. Then the set inclusion*

$$\{C^-\} \subseteq \{\, A_1^- + A_2^- + \cdots + A_k^- \,\} \tag{23.40}$$

holds if and only if

$$R(A_1)\cap R(A_2)\cap\cdots\cap R(A_k) = \{0\}, \quad or \quad R(A_1^T)\cap R(A_2^T)\cap\cdots\cap R(A_k^T) = \{0\}, \tag{23.41}$$

or

$$r(\,N - QCP\,) = r\begin{bmatrix} N \\ P \end{bmatrix} + r[\,N,\ Q\,] - r(N) + r(C) - m - n, \tag{23.42}$$

where $N,\ P$ and Q are defined in Theorem 23.21.

The result in Theorem 23.26 implies the following special case.

Corollary 23.27. *Let $A_1,\ A_2,\ \cdots,\ A_k \in \mathcal{F}^{m\times n}$ be given. Then the equality*

$$\{\, A_1^- + A_2^- + \cdots + A_k^- \,\} = \{0^-\} = \mathcal{F}^{n\times m} \tag{23.43}$$

holds if and only if $A_1,\ A_2,\ \cdots,\ A_k$ satisfy (23.41).

Theorem 23.28. *Let $A_1,\ A_2,\ \cdots,\ A_k,\ C \in \mathcal{F}^{m\times n}$ be given with*

$$\cap_{i=1}^k R(A_i) \neq \{0\} \quad and \quad \cap_{i=1}^k R(A_i^T) \neq \{0\}. \tag{23.44}$$

Then the equality

$$\{\, A_1^- + A_2^- + \cdots + A_k^- \,\} = \{C^-\} \tag{23.45}$$

holds if and only if they satisfy the following rank additivity condition

$$r\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix} = r\begin{bmatrix} N \\ P \end{bmatrix} + r(Q) = r[\,N,\ Q\,] + r(P), \tag{23.46}$$

and

$$C = -[\,0,\ I_m\,]\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}^- \begin{bmatrix} 0 \\ I_n \end{bmatrix}, \tag{23.47}$$

where $N,\ P$ and Q are defined in Theorem 23.21. In that case, the rank of C satisfies the equality

$$r(C) = r(N) + r(Q) - r[\,N,\ Q\,] = r(N) + r(P) - r\begin{bmatrix} N \\ P \end{bmatrix}, \tag{23.48}$$

or more precisely

$$r(C) = \dim[\,\cap_{i=1}^k R(A_i)\,] = \dim[\,\cap_{i=1}^k R(A_i^T)\,]. \tag{23.49}$$

Proof. Assume first that (23.45) holds. Then it follows from Theorem 23.26 and (23.44) that $A_1,\ A_2,\ \cdots,\ A_k$, and C satisfy (23.42). On the other hand, (23.45) implies that

$$\min_{A_1^-,\,\cdots,\,A_k^-} r(\,A_1^- + \cdots + A_k^-\,) = \min_{C^-} r(C^-), \tag{23.50}$$

which, by the first equality in Theorem 23.24, is equivalent to

$$r(C) = r\begin{bmatrix} N & Q \\ Q & 0 \end{bmatrix} - r\begin{bmatrix} N \\ P \end{bmatrix} - r[\,N,\ Q\,] + r(N). \tag{23.51}$$

Putting it (23.42) yields

$$r(\,N - QCP\,) = r\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix} - m - n. \tag{23.52}$$

On the other hand, it is easy to verify that

$$r\begin{bmatrix} N & Q & 0 \\ P & 0 & I_n \\ 0 & I_m & -C \end{bmatrix} = r\begin{bmatrix} N - QCP & 0 & 0 \\ 0 & 0 & I_n \\ 0 & I_m & 0 \end{bmatrix} = m + n + r(\,N - QCP\,).$$

Thus (23.52) is equivalent to

$$r\begin{bmatrix} N & Q & 0 \\ P & 0 & I_n \\ 0 & I_m & -C \end{bmatrix} = r\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}.$$

In light of (1.5), the rank equality is further equivalent to

$$r\begin{bmatrix} N & Q & 0 \\ P & 0 & I_n \end{bmatrix} = r\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}, \quad r\begin{bmatrix} N & Q \\ P & 0 \\ 0 & I_m \end{bmatrix} = r\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}, \tag{23.53}$$

and

$$C = -[\,0, \; I_m\,]\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}^{-}\begin{bmatrix} 0 \\ I_n \end{bmatrix},$$

which are exactly (23.46) and (23.47). Consequently combining (23.51) with (23.46) yields (23.48), and then yields (23.49) by Lemma 23.1. Conversely if (23.46)—(23.48) hold, then it is easy to verify that (23.42) and Theorem 23.23(c) are all satisfied, both of which imply that (23.45) holds. □

On the basis of Theorem 23.28, we now can reasonably extend the concept of parallel sums of two matrices to k matrices.

Defination. The k matrices $A_1, \, A_2, \, \cdots, \, A_k \in \mathcal{F}^{m\times n}$ are said to be parallel summable, if the matrix product

$$-[\,0, \; I_m\,]\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}^{-}\begin{bmatrix} 0 \\ I_n \end{bmatrix} \tag{23.54}$$

is invariant with respect to the choice of the inner inverse in it, where $N, \; P$ and Q are defined in Theorem 23.21. In that case, the matrix product in (23.54) is called the parallel sum of $A_1, \, A_2, \, \cdots, \, A_k$ and denoted by $p(\,A_1, \, A_2, \, \cdots, \, A_k\,)$.

Various properties on parallel sums of k matrices can easily be derived from Theorem 23.28. Below are some of them, which are quite analogous to those for parallel sums of two matrices.

Theorem 23.29. *A null matrix is parallel summable with any other matrices of the same size, and their parallel sum is also a null matrix.*

Proof. Let $A = \begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}$ in (23.54). If one of $A_1, \, A_2, \, \cdots, \, A_k$ is null, then it is easy to verify that

$$r\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix} = r(N) + m + n.$$

In that case, applying Corollary 21.7(a) to (23.54), we obtain

$$\begin{aligned} \max_{A^-} r\left([\,0, \; I_m\,]A^{-}\begin{bmatrix} 0 \\ I_n \end{bmatrix} \right) &= \min\left\{ m, \quad n, \quad r\begin{bmatrix} N & Q & 0 \\ P & 0 & I_n \\ 0 & I_m & 0 \end{bmatrix} - r\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix} \right\} \\ &= \min\left\{ m, \quad n, \quad m + n + r(N) - \begin{bmatrix} N & Q \\ P & 0 \end{bmatrix} \right\} = 0. \end{aligned}$$

This result implies that (23.54) is always null with respect to the choice of A^-. Thus $A_1, \, A_2, \, \cdots, \, A_k$ are parallel summable. □

Theorem 23.30. *Let $A_1, \, A_2, \, \cdots, \, A_k \in \mathcal{F}^{m\times n}$ be nonnull matrices. Then they are parallel summable if and only if*

$$R\begin{bmatrix} 0 \\ I_n \end{bmatrix} \subseteq R\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix} \quad \textit{and} \quad R(\,[\,0, \; I_m\,]^T\,) \subseteq R\left(\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}^T \right), \tag{23.55}$$

or equivalently

$$r\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix} = r\begin{bmatrix} N \\ P \end{bmatrix} + r(Q) = r[\,N, \; Q\,] + r(P), \tag{23.56}$$

where $N,\ P$ and Q are defined in Theorem 23.1.

Proof. It is well-known that (see [118], [103]) that a product AB^-C is invariant with respect to the choice of A^- if and only if $R(A^T) \subseteq R(B^T)$ and $R(C) \subseteq R(B)$. Applying this assertion to (23.54) immediately leads to (23.55). The equivalence of (23.55) and (23.56) is obvious. □

Theorem 23.31. *Let $A_1,\ A_2,\ \cdots,\ A_k \in \mathcal{F}^{m\times n}$ be given. If they are parallel summable, then*

(a) $\{\, [\, p(\, A_1,\ A_2,\ \cdots,\ A_k\,)\,]^- \,\} = \{\, A_1^- + A_2^- + \cdots + A_k^- \,\}$.

(b) $p(\, A_1,\ A_2,\ \cdots,\ A_k\,) = p(\, A_{i_1},\ A_{i_2},\ \cdots,\ A_{i_k}\,)$, *where* $i_1,\ i_2,\ \cdots,\ i_k$ *are any permutation of* $1,\ 2,\ \cdots,\ k$.

(c) $p(\, A_1^T,\ A_2^T,\ \cdots,\ A_k^T\,) = [\, p(\, A_1,\ A_2,\ \cdots,\ A_k\,)\,]^T$.

Proof. If any one of $A_1,\ A_2,\ \cdots,\ A_k$ is null, then Parts (a)—(c) are naturally valid by Theorem 23.29. Now suppose that $A_1,\ A_2,\ \cdots,\ A_k$ are nonnull and parallel summable. Then by Theorems 23.30 and 23.28, we immediately see that the equality in Part (a) holds. The equality in Part (b) comes from a trivial equality $\{\, A_1^- + A_2^- + \cdots + A_k^- \,\} = \{\, A_{i_1}^- + A_{i_2}^- + \cdots + A_{i_k}^- \,\}$, and Theorem 21.10(c). By Theorem 23.30, we also know that if $A_1,\ A_2,,\ \cdots,\ A_k$ satisfy (23.56), then $A_1^T,\ A_2^T,\ \cdots,\ A_k^T$ naturally satisfy

$$r\begin{bmatrix} N^T & P^T \\ Q^T & 0 \end{bmatrix} = r\begin{bmatrix} N^T \\ Q^T \end{bmatrix} + r(P^T) = r[\, N^T,\ P^T\,] + r(Q^T),$$

where $N,\ P$ and Q are defined in Theorem 23.21. Thus $A_1^T,\ A_2^T,\ \cdots,\ A_k^T$ are also parallel summable. In that case, it follows from (23.54) that

$$\begin{aligned} p(\, A_1^T,\ A_2^T,\ \cdots,\ A_k^T\,) &= -[\, 0,\ I_n\,]\begin{bmatrix} N^T & P^T \\ Q^T & 0 \end{bmatrix}^- \begin{bmatrix} 0 \\ I_m \end{bmatrix} \\ &= -[\, 0,\ I_n\,]\left(\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}^-\right)^T \begin{bmatrix} 0 \\ I_m \end{bmatrix} \\ &= -\left([\, 0,\ I_m\,]\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}^- \begin{bmatrix} 0 \\ I_n \end{bmatrix}\right)^T = [\, p(\, A_1,\ A_2,\ \cdots,\ A_k\,)\,]^T, \end{aligned}$$

which is the result in Part (c). □

Theorem 23.32. *Let $A_1,\ A_2,\ \cdots,\ A_k \in \mathcal{F}^{m\times n}$ be given, and $B \in \mathcal{F}^{m\times m}$, $C \in \mathcal{F}^{n\times n}$ are two nonsingular matrices. Then $A_1,\ A_2,\ \cdots,\ A_k$ are parallel summable if and only if $BA_1C,\ BA_2C,\ \cdots,\ BA_kC$ are are parallel summable. In that case,*

$$p(\, BA_1C,\ BA_2C,\ \cdots,\ BA_kC\,) = Bp(\, A_1,\ A_2,\ \cdots,\ A_k\,)C. \tag{23.57}$$

Proof. If any one of $A_1,\ A_2,\ \cdots,\ A_k$ is null, then (23.57) is a trivial result by Theorem 23.30. Now suppose that $A_1,\ A_2,\ \cdots,\ A_k$ are nonnull and denote

$$\widehat{B} = \mathrm{diag}(\, B,\ B,\ \cdots,\ B\,), \quad \widehat{C} = \mathrm{diag}(\, C,\ C,\ \cdots,\ C\,).$$

Since B and C are nonsingular, $\widehat{B}$ and $\widehat{C}$ are nonsingular, too. In that case, it is easy to verify that

$$\begin{aligned} & r\begin{bmatrix} \widehat{B}N\widehat{C} & Q \\ P & 0 \end{bmatrix} = r\begin{bmatrix} N & \widehat{B}^{-1}Q \\ P\widehat{C}^{-1} & 0 \end{bmatrix} = r\begin{bmatrix} N & \widehat{B}^{-1}QB \\ CP\widehat{C}^{-1} & 0 \end{bmatrix} = r\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}, \\ & r\begin{bmatrix} \widehat{B}N\widehat{C} \\ P \end{bmatrix} = r\begin{bmatrix} N \\ P\widehat{C}^{-1} \end{bmatrix} = r\begin{bmatrix} N \\ CP\widehat{C}^{-1} \end{bmatrix} = r\begin{bmatrix} N \\ P \end{bmatrix}, \\ & r[\, \widehat{B}N\widehat{C},\ Q\,] = r[\, N,\ \widehat{B}^{-1}Q\,] = r[\, N,\ \widehat{B}^{-1}QB\,] = r[\, N,\ Q\,]. \end{aligned}$$

Combining them with (23.56) clearly shows that $A_1,\ A_2,\ \cdots,\ A_k$ are parallel summable if and only if $BA_1C,\ BA_2C,\ \cdots,\ BA_kC$ are parallel summable. From the nonsingularity of B and C, we also see that

$$\begin{bmatrix} \widehat{B}N\widehat{C} & Q \\ P & 0 \end{bmatrix}^- = \begin{bmatrix} \widehat{C}^{-1} & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}^- \begin{bmatrix} \widehat{B}^{-1} & 0 \\ 0 & C \end{bmatrix}.$$

Thus it follows from (23.54) that

$$\begin{aligned} p(\,BA_1C,\ BA_2C,\,\cdots,\,BA_kC\,) &= -[\,0,\ I_m\,]\begin{bmatrix} \widehat{B}N\widehat{C} & Q \\ P & 0 \end{bmatrix}^{-}\begin{bmatrix} 0 \\ I_n \end{bmatrix} \\ &= -[\,0,\ I_m\,]\begin{bmatrix} \widehat{C}^{-1} & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}^{-}\begin{bmatrix} \widehat{B}^{-1} & 0 \\ 0 & C \end{bmatrix}\begin{bmatrix} 0 \\ I_n \end{bmatrix} \\ &= -B[\,0,\ I_m\,]\begin{bmatrix} N & Q \\ P & 0 \end{bmatrix}^{-}\begin{bmatrix} 0 \\ I_n \end{bmatrix}C \\ &= Bp(\,A_1,\,A_2,\,\cdots,\,A_k\,)C, \end{aligned}$$

which is (23.57). $\square$

Chapter 24

Ranks and independence of submatrices in solutions to $BXC = A$

Suppose that $BXC = A$ is a consistent matrix equation over an arbitrary field $\mathcal{F}$, where $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ and $C \in \mathcal{F}^{l\times n}$ are given. Then it can factor in the form

$$[\,B_1,\ B_2\,]\begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix}\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = A, \tag{24.1}$$

where $X_1 \in \mathcal{F}^{k_1\times l_1}$, $X_2 \in \mathcal{F}^{k_1\times l_2}$, $X_3 \in \mathcal{F}^{k_2\times l_1}$ and $X_4 \in \mathcal{F}^{k_2\times l_2}$, $k_1 + k_2 = k$, $l_1 + l_2 = l$. In this chapter, we determine maximal and minimal possible ranks of submatrices X_1—X_4 in a solution to (24.1).

Possible ranks of solutions of linear matrix equations and various related topics have been considered previously by several authors. For example, Mitra in [98] examined solutions with fixed ranks for the matrix equations $AX = B$ and $AXB = C$; Mitra in [99] gave common solutions of minimal rank of the pair of matrix equations $AX = C$, $XB = D$; Uhlig in [142] presented maximal and minimal possible ranks of solutions of the equation $AX = B$; Mitra [103] described common solutions with the minimal rank to the pair of matrix equations $A_1XB_1 = C_1$ and $A_2XB_2 = C_2$. Besides the work in the chapter, we shall also consider in the next two chapters possible ranks of the two real matrices X_0 and X_1 in solutions to the complex matrix equation $B(\,X_0 + iX_1\,)C = A$, as well as possible ranks and independence of solutions to the matrix equation $B_1XC_1 + B_2YC_2 = A$.

For convenience of representation, we adopt the notation for the collections of the submatrices X_1—X_4 in (24.1)

$$S_i = \left\{ X_i \;\middle|\; [\,B_1,\ B_2\,]\begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix}\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = A \right\}, \qquad i = 1,\, 2,\, 3,\, 4. \tag{24.2}$$

It is easily seen that X_1—X_4 in (24.1) can be written as

$$X_1 = [\,I_{k_1},\ 0\,]X\begin{bmatrix} I_{l_1} \\ 0 \end{bmatrix} = P_1XQ_1, \qquad X_2 = [\,I_{k_1},\ 0\,]X\begin{bmatrix} 0 \\ I_{l_2} \end{bmatrix} = P_1XQ_2, \tag{24.3}$$

$$X_3 = [\,0,\ I_{k_2}\,]X\begin{bmatrix} I_{l_1} \\ 0 \end{bmatrix} = P_2XQ_1, \qquad X_4 = [\,0,\ I_{k_2}\,]X\begin{bmatrix} 0 \\ I_{l_2} \end{bmatrix} = P_2XQ_2. \tag{24.4}$$

Since $BXC = A$ is consistent, its general solution can be written as $X = B^-AC^- + F_BV + WE_C$. Putting it in (24.3) and (24.4) yields the general expressions of X_1—X_4 as follows

$$X_1 = P_1X_0Q_1 + P_1F_BV_1 + W_1E_CQ_1, \quad X_2 = P_1X_0Q_2 + P_1F_BV_2 + W_1E_CQ_2, \tag{24.5}$$

$$X_3 = P_2X_0Q_1 + P_2F_BV_1 + W_2E_CQ_1, \quad X_4 = P_2X_0Q_2 + P_2F_BV_2 + W_2E_CQ_2, \tag{24.6}$$

where $X_0 = B^-AC^-,\ V = [\,V_1,\ V_2\,]$ and $W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}$.

Theorem 24.1. *Suppose that the matrix equation* (24.1) *is consistent. Then*

$$\max_{X_1 \in S_1} r(X_1) = \min\left\{ k_1, \quad l_1, \quad r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} - r(B) - r(C) + k_1 + l_1 \right\}, \tag{24.7}$$

$$\min_{X_1 \in S_1} r(X_1) = r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} - r(B_2) - r(C_2). \tag{24.8}$$

Proof. It is quite obvious that to determine maximal and minimal ranks of X_1 in (24.1) is in fact to determine maximal and minimal ranks of P_1XQ_1 subject to the consistent equation $BXC = A$. Thus applying (20.3) and (20.4) to $X_1 = P_1XQ_1$ produces the following two expressions

$$\max_{BXC=A} r(P_1XQ_1) = \min\left\{ r(P_1), \quad r(Q_1), \quad r\begin{bmatrix} 0 & 0 & P_1 \\ 0 & A & B \\ Q_1 & C & 0 \end{bmatrix} - r(B) - r(C) \right\},$$

$$\min_{BXC=A} r(P_1XQ_1) = r\begin{bmatrix} 0 & 0 & P_1 \\ 0 & A & B \\ Q_1 & C & 0 \end{bmatrix} - r\begin{bmatrix} P_1 \\ B \end{bmatrix} - r[\, Q_1, \ C\,].$$

Putting the given matrices $B = [\, B_1, \ B_2\,]$, $C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$, P_1 and Q_1 in them and simplifying yields the desired formulas (24.7) and (24.8). The details are omitted. $\square$

Maximal and minimal ranks of the submatrices X_2, X_3, and X_4 in (24.1) can also be derived in the same manner. We omit them here for simplicity. The two formulas in (24.7) and (24.8) can help to characterize structure of solutions to (24.1). Next are some of them.

Corollary 24.2. *Suppose that the matrix equation* (24.1) *is consistent. Then*

(a) *Eq.* (24.1) *has a solution with the form* $X = \begin{bmatrix} 0 & X_2 \\ X_3 & X_4 \end{bmatrix}$ *if and only if* $r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} = r(B_2) + r(C_2)$.

(b) *All the solutions of* (24.1) *have the form* $X = \begin{bmatrix} 0 & X_2 \\ X_3 & X_4 \end{bmatrix}$ *if and only if*

$$\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} = r(B) + r(C) - k_1 - l_1, \tag{24.9}$$

or equivalently

$$r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} = r(B_2) + r(C_2), \ r(B_1) = k_1, \ \ r(C_1) = l_1, \ \ R(B_1) \cap R(B_2) = \{0\} \ \textit{and} \ \ R(C_1^T) \cap R(C_2^T) = \{0\}. \tag{24.10}$$

Proof. Part (a) and (24.9) follows directly from (24.7) and (24.8). On the other hand, observe that

$$\begin{aligned} & r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} - r(B) - r(C) + k_1 + l_1 \\ = & \left(r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} - r(B_2) - r(C_2) \right) + [\, k_1 + r(B_2) - r(B)\,] + [\, l_1 + r(C_2) - r(C)\,]. \end{aligned}$$

Thus (24.9) is equivalent to (24.10). $\square$

Theorem 24.3. *Suppose that the matrix equation* (24.1) *is consistent. Then*

(a) *Eq.* (24.1) *has a solution with the form* $X = \begin{bmatrix} X_1 & 0 \\ X_3 & 0 \end{bmatrix}$ *if and only if* $R(A^T) \subseteq R(C_1^T)$.

(b) *Eq.* (24.1) *has a solution with the form* $X = \begin{bmatrix} X_1 & X_2 \\ 0 & 0 \end{bmatrix}$ *if and only if* $R(A) \subseteq R(B_1)$.

(c) *Eq.* (24.1) *has a solution with the form* $X = \begin{bmatrix} X_1 & 0 \\ 0 & 0 \end{bmatrix}$ *if and only if* $R(A) \subseteq R(B_1)$ *and* $R(A^T) \subseteq R(C_1^T)$.

Proof. According to (20.3) and (20.4), we find that

$$\min_{BXC=A} r\begin{bmatrix} X_2 \\ X_4 \end{bmatrix} = \min_{BXC=A} r(XQ_2) = r\begin{bmatrix} A \\ C_1 \end{bmatrix} - r(C_1),$$
$$\min_{BXC=A} r[\, X_3, \ X_4 \,] = \min_{BXC=A} r(P_2X) = r[\, A, \ B_1 \,] - r(B_1).$$

Thus we have Parts (a) and (b). The result in Part (c) is evident. □

Note from (24.5) and (24.6) that X_1 and X_4, X_2 and X_3 are independent in their expressions, i.e., both of them do not involve the same variant matrices, thus we have the following.

Theorem 24.4. *Suppose that the matrix equation* (24.1) *is consistent. Then*

(a) *Eq.* (24.1) *must have two solutions with the forms*

$$X = \begin{bmatrix} \widehat{X_1} & X_2 \\ X_3 & \widehat{X_4} \end{bmatrix}, \qquad X = \begin{bmatrix} X_1 & \widehat{X_2} \\ \widehat{X_3} & X_4 \end{bmatrix},$$

where $\widehat{X_1}$—$\widehat{X_4}$ *in them with the ranks*

$$r(\widehat{X_1}) = \min_{X_1 \in S_1} r(X_1) = r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} - r(B_2) - r(C_2),$$
$$r(\widehat{X_2}) = \min_{X_2 \in S_2} r(X_2) = r\begin{bmatrix} A & B_2 \\ C_1 & 0 \end{bmatrix} - r(B_2) - r(C_1),$$
$$r(\widehat{X_3}) = \min_{X_3 \in S_3} r(X_3) = r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} - r(B_1) - r(C_2),$$
$$r(\widehat{X_4}) = \min_{X_4 \in S_4} r(X_4) = r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} - r(B_1) - r(C_1).$$

(b) *Eq.* (24.1) *has a solution with the form* $X = \begin{bmatrix} 0 & X_2 \\ X_3 & 0 \end{bmatrix}$, *if and only if*

$$r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} = r(B_1) + r(C_1), \quad \textit{and} \quad r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} = r(B_2) + r(C_2).$$

(c) *Eq.* (24.1) *has a solution with the form* $X = \begin{bmatrix} X_1 & 0 \\ 0 & X_4 \end{bmatrix}$, *if and only if*

$$r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} = r(B_1) + r(C_2), \quad \textit{and} \quad r\begin{bmatrix} A & B_2 \\ C_1 & 0 \end{bmatrix} = r(B_2) + r(C_1).$$

The result in Theorem 24.4(c) in fact implies a necessary and sufficient condition for the matrix equation $B_1X_1C_1 + B_2X_4C_2 = A$ to be solvable, which was first established by Özgüler in [110].

The uniqueness of the submatrices X_1—X_4 in (24.1) can be determined by (24.5) and (24.6).

Theorem 24.5. *Suppose that the matrix equation* (24.1) *is consistent. The submatrix* X_1 *in* (24.1) *is unique if and only if* (24.1) *satisfies the following four conditions*

$$r(B_1) = k_1, \quad r(C_1) = l_1, \quad R(B_1) \cap R(B_2) = \{0\}, \quad R(C_1^T) \cap R(C_2^T) = \{0\}. \tag{24.11}$$

Proof. It is easy to see from (24.5) that X_1 is unique if and only if $P_{,1}F_B = 0$ and $E_CQ_1 = 0$, where we find by (1.2) and (1.3) that

$$P_1F_B = 0 \Rightarrow r\begin{bmatrix} P_1 \\ B \end{bmatrix} = r(B) \Rightarrow k_1 + r(B_2) = r(B) \Rightarrow r(B_1) = k_1 \text{ and } R(B_1) \cap R(B_2) = \{0\},$$
$$E_CQ_1 = 0 \Rightarrow r[\, Q_1, \ C \,] = r(C) \Rightarrow l_1 + r(C_2) = r(C) \Rightarrow r(C_1) = l_1 \text{ and } R(C_1^T) \cap R(C_2^T) = \{0\}.$$

Thus we have (24.11). □

The following result is concerning the independence of submatrices in solutions to (24.1).

Theorem 24.6. *Suppose that the matrix equation* (24.1) *is consistent with* $B \neq 0$ *and* $C \neq 0$.

(a) *Consider* S_1—S_4 *in* (24.1) *as four independent matrix sets. Then*

$$\begin{aligned}&\max_{X_i \in S_i} r\left(A - [\, B_1, \ B_2\,] \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \right) \\ &= \min\{r(B), \quad r(C), \quad r(B_1) + r(B_2) - r(B) + r(C_1) + r(C_2) - r(C)\}.\end{aligned} \tag{24.12}$$

(b) *The four submatrices* X_1—X_4 *in* (24.1) *are independent, that is, for any choice of* $X_i \in S_i (i = 1, 2, 3, 4)$, *the corresponding matrix* $X = \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix}$ *is a solution of* (24.1), *if and only if*

$$R(B_1) \cap R(B_2) = \{0\} \quad and \quad R(C_1^T) \cap R(C_2^T) = \{0\}. \tag{24.13}$$

Proof. According to (24.5) and (24.6), the general expressions of X_1—X_4 in S_1—S_4 can independently be written as

$$X_1 = P_1X_0Q_1 + P_1F_BV_1 + W_1E_CQ_1, \quad X_2 = P_1X_0Q_2 + P_1F_BV_2 + W_2E_CQ_2,$$

$$X_3 = P_2X_0Q_1 + P_2F_BV_3 + W_3E_CQ_1, \quad X_4 = P_2X_0Q_2 + P_2F_BV_4 + W_4E_CQ_2,$$

where $X_0 = B^-AC^-$, V_1—V_4 and W_1—W_4 are arbitrary. Putting them in X yields

$$\begin{aligned}&\begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} \\ &= \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} X_0[\, Q_1, \ Q_2\,] + \begin{bmatrix} P_1F_B & 0 \\ 0 & P_2F_B \end{bmatrix} \begin{bmatrix} V_1 & V_2 \\ V_3 & V_4 \end{bmatrix} + \begin{bmatrix} W_1 & W_2 \\ W_3 & W_4 \end{bmatrix} \begin{bmatrix} E_CQ_1 & 0 \\ 0 & E_CQ_2 \end{bmatrix} \\ &= X_0 + GV + WH,\end{aligned}$$

where $G = \mathrm{diag}(\, P_1F_B, \ P_2F_B\,)$, $H = \mathrm{diag}(\, E_CQ_1, \ E_CQ_2\,)$. Applying (1.6) to it, we find

$$\begin{aligned}&\max_{X_i \in S_i} r\left(A - [\, B_1, \ B_2\,] \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \right) \\ &= \max_{V, W} r(\, BGVC + BWHC\,) = \min\{\, r(B), \quad r(C), \quad r(BG) + r(HC)\,\}.\end{aligned} \tag{24.14}$$

According to (1.2) and (1.3), we see that

$$\begin{aligned} r(BG) &= r[\, B_1P_1F_B, \ B_2P_2F_B\,] \\ &= r\begin{bmatrix} B_1P_1 & B_2P_2 \\ B & 0 \\ 0 & B \end{bmatrix} - 2r(B) = r\begin{bmatrix} B_1 & 0 & 0 & B_2 \\ B_1 & B_2 & 0 & 0 \\ 0 & 0 & B_1 & B_2 \end{bmatrix} - 2r(B) = r(B_1) + r(B_2) - r(B),\end{aligned}$$

$$\begin{aligned}&r(HC) \\ &= r\begin{bmatrix} E_CQ_1C_1 \\ E_CQ_2C_2 \end{bmatrix} = r\begin{bmatrix} Q_1C_1 & C & 0 \\ Q_2C_2 & 0 & C \end{bmatrix} - 2r(C) = r\begin{bmatrix} C_1 & C_1 & 0 \\ 0 & C_2 & 0 \\ 0 & 0 & C_1 \\ C_2 & 0 & C_2 \end{bmatrix} - 2r(C) = r(C_1) + r(C_2) - r(C).\end{aligned}$$

Putting them in (24.14), we obtain (24.12). The result in Part (b) is a direct consequence of (24.12). □

Let

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{24.15}$$

be a partitioned matrix over $\mathcal{C}$, where $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$, $C \in \mathcal{F}^{l\times n}$ and $D \in \mathcal{F}^{l\times k}$, and write its inner inverse in the block form

$$M^- = \begin{bmatrix} G_1 & G_2 \\ G_3 & G_4 \end{bmatrix}, \tag{24.16}$$

where $G_1 \in \mathcal{F}^{n\times m}$. In this section, we determine maximal and minimal ranks of the blocks G_1—G_4 in (24.16) and consider their relationship with A, B, C and D.

For convenience of representation, we adopt the notation

$$T_i = \left\{ G_i \;\middle|\; \begin{bmatrix} G_1 & G_2 \\ G_3 & G_4 \end{bmatrix} \in \{M^-\} \right\}, \qquad i = 1, 2, 3, 4. \tag{24.17}$$

Notice that M^- is in fact a solution to the matrix equation $MXM = M$. Thus applying the results in Theorem 24.1 to (24.15) and (24.16), we find the following.

Theorem 24.7. *Let M and M^- be given by* (24.15) *and* (24.16). *Then*

$$\max_{G_1\in T_1} r(G_1) = \min\{\, m, \quad n, \quad m + n + r(D) - r(M)\,\}, \tag{24.18}$$

$$\min_{G_1\in T_1} r(G_1) = r(M) + r(D) - r[\,C,\ D\,] - r\begin{bmatrix} B \\ D \end{bmatrix}. \tag{24.19}$$

Proof. Follows from (24.7) and (24.8). □

Corollary 24.8. *Let M and M^- be given by* (24.15) *and* (24.16). *Then*

(a) *M has a g-inverse with the form $M^- = \begin{bmatrix} 0 & G_2 \\ G_3 & G_4 \end{bmatrix}$ if and only if $r(M) = r\begin{bmatrix} B \\ D \end{bmatrix} + r[\,C,\ D\,] - r(D)$.*

(b) *All the g-inverses of M have the form $M^- = \begin{bmatrix} 0 & G_2 \\ G_3 & G_4 \end{bmatrix}$ if and only if $r(M) = m + n - r(D)$.*

Proof. Follows from Theorem 24.7. □

Corollary 24.9. *Let M and M^- be given by* (24.15) *and* (24.16).

(a) *M has a g-inverse with the form $M^- = \begin{bmatrix} G_1 & 0 \\ G_3 & 0 \end{bmatrix}$ if and only if $R([\,C,\ D\,]^T) \subseteq R([\,A,\ B\,]^T)$.*

(b) *M has a g-inverse with the form $M^- = \begin{bmatrix} G_1 & G_2 \\ 0 & 0 \end{bmatrix}$ if and only if $R\begin{bmatrix} B \\ D \end{bmatrix} \subseteq R\begin{bmatrix} A \\ C \end{bmatrix}$.*

(c) *M has a g-inverse with the form $M^- = \begin{bmatrix} G_1 & 0 \\ 0 & 0 \end{bmatrix}$ if and only if $r(M) = r(A)$.*

Proof. Follows from Corollary 24.3. □

Corollary 24.10. *Let M and M^- be given by* (24.15) *and* (24.16). *Then*

(a) *M has a g-inverse with the form $M^- = \begin{bmatrix} G_1 & 0 \\ 0 & G_4 \end{bmatrix}$, if and only if*

$$r(M) = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,C,\ D\,] - r(C) = r\begin{bmatrix} B \\ D \end{bmatrix} + r[\,A,\ B\,] - r(B).$$

(b) *M has a g-inverse with the form $M^- = \begin{bmatrix} 0 & G_2 \\ G_3 & 0 \end{bmatrix}$, if and only if*

$$r(M) = r\begin{bmatrix} B \\ D \end{bmatrix} + r[\,C,\ D\,] - r(D) = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,] - r(A).$$

Proof. Follows from Theorem 24.4 (b) and (c). □

Corollary 24.11 (Rao and Yanai [119]). *Let M and M^- be given by* (24.15) *and* (24.16). *Then the submatrix G_1 in* (24.16) *is unique if and only if M satisfies the following three conditions*

$$r[\,A,\ B\,] = m, \quad r\begin{bmatrix} A \\ C \end{bmatrix} = n, \quad r(M) = n + r\begin{bmatrix} B \\ D \end{bmatrix} = m + r[\,C,\ D\,].$$

Proof. Follows from Theorem 24.5. □

Theorem 24.12. *Let M and M^- be given by* (24.15) *and* (24.16).

(a) *Consider T_1—T_4 in* (24.17) *as four independent matrix sets. Then*

$$\max_{G_i \in T_i} r\left(M - M\begin{bmatrix} G_1 & G_2 \\ G_3 & G_4 \end{bmatrix} M \right) = \min\left\{ r(M), \quad r\begin{bmatrix} A \\ C \end{bmatrix} + r\begin{bmatrix} B \\ D \end{bmatrix} + r[\,A,\ B\,] + r[\,C,\ D\,] - 2r(M) \right\}.$$

(b) (Rao and Yanai [119]) *The four submatrices G_1—G_4 in* (24.16) *are independent if and only if M satisfies the following rank additivity condition*

$$r(M) = r\begin{bmatrix} A \\ C \end{bmatrix} + r\begin{bmatrix} B \\ D \end{bmatrix} = r[\,A,\ B\,] + r[\,C,\ D\,].$$

Proof. Follows from Theorem 24.6. □

In the remainder of this section, we consider the relationship between $\{A^-\}$ and T_1, $\{B^-\}$ and T_3, $\{C^-\}$ and T_2, $\{D^-\}$ and T_4, where T_1—T_4 are defined in (24.17).

Theorem 24.13. *Let M and M^- be given by* (24.15) *and* (24.16). *Then*

$$\max_{G_1 \in T_1} r(\,A - AG_1A\,) = \min\left\{ r(A), \quad r(A) + r\begin{bmatrix} 0 & B \\ C & D \end{bmatrix} - r(M) \right\}, \tag{24.20}$$

$$\min_{G_1 \in T_1} r(\,A - AG_1A\,) = r(A) + r(M) + r\begin{bmatrix} 0 & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix}. \tag{24.21}$$

Proof. Let $P = [\,I_n,\ 0\,]$ and $Q = [\,I_m,\ 0\,]^T$. Then according to (22.1) and (22.2), we find that

$$\begin{aligned} \max_{G_1 \in T_1} r(\,A - AG_1A\,) &= \max_{M^-} r(\,A - APM^-QA\,) \\ &= \min\left\{ r(AP), \quad r(QA), \quad r\begin{bmatrix} M & QA \\ AP & A \end{bmatrix} - r(M) \right\} \\ &= \min\{\, r(A), \quad r(\,M - QAP\,) + r(A) - r(M)\,\} \\ &= \min\left\{ r(A), \quad r\begin{bmatrix} 0 & B \\ C & D \end{bmatrix} + r(A) - r(M) \right\}, \end{aligned}$$

$$\begin{aligned} \min_{G_1 \in T_1} r(\,A - AG_1A\,) &= \min_{M^-} r(\,A - APM^-QA\,) \\ &= r(M) - r[\,M,\ QA\,] - r\begin{bmatrix} M \\ AP \end{bmatrix} + r\begin{bmatrix} M & QA \\ AP & A \end{bmatrix} \\ &= r(A) + r(M) + r(\,M - QAP\,) - r[\,M,\ QA\,] - r\begin{bmatrix} M \\ AP \end{bmatrix} \\ &= r(A) + r(M) + r\begin{bmatrix} 0 & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix}, \end{aligned}$$

establishing (24.20) and (24.21). □

A similar result to (24.21) was presented in (21.104).

Corollary 24.14. *Let M and M^- be given by* (24.15) *and* (24.16). *Then*

(a) *M has a g-inverse with the form* $M^- = \begin{bmatrix} A^- & G_2 \\ G_3 & G_4 \end{bmatrix}$ *if and only if*

$$r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix} + r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} = r(A) + r(M) + r\begin{bmatrix} 0 & B \\ C & D \end{bmatrix}.$$

(b) $T_1 \subseteq \{A^-\}$, *i.e., any* G_1 *in* T_1 *is a g-inverse of* A *if and only if* $r(M) = r(A) + r\begin{bmatrix} 0 & B \\ C & D \end{bmatrix}$.

Proof. Follows immediately from Theorem 24.13. □

Corollary 24.15. *Let* M *and* M^- *be given by* (24.15) *and* (24.16), *and* T_1—T_4 *are given by* (24.17). *Then*

$$T_1 \subseteq \{A^-\}, \qquad T_2 \subseteq \{C^-\}, \qquad T_3 \subseteq \{B^-\}, \qquad T_4 \subseteq \{D^-\} \tag{24.22}$$

are all satisfied if and only if

$$r(M) = r(A) + r(B) + r(C) + r(D). \tag{24.23}$$

Proof. If (24.22) holds, then

$$\begin{aligned} r(M) = r(MM^-) = tr(MM^-) &= tr\begin{bmatrix} AG_1 + BG_3 & AG_2 + BG_4 \\ CG_1 + DG_3 & CG_2 + DG_4 \end{bmatrix} \\ &= tr(AG_1) + tr(BG_3) + tr(CG_2) + tr(DG_4) \\ &= tr(AA^-) + tr(BB^-) + tr(CC^-) + tr(DD^-) \\ &= r(A) + r(B) + r(C) + r(D). \end{aligned}$$

Conversely, if (24.23) is satisfied, then (24.22) naturally holds by Corollary 24.14(b). □

When $D = 0$ in the above theorems and corollaries, the corresponding results can further simplify. We leave them to the reader.

Chapter 25

Extreme ranks of X_0 and X_1 in solutions to $B(X_0 + iX_1)C = A$

Suppose $BXC = A$ is a complex matrix equation. Then it can be written as

$$(B_0 + iB_1)(X_0 + iX_1)(C_0 + iC_1) = (A_0 + iA_1), \tag{25.1}$$

where $A_0,\ A_1 \in \mathcal{R}^{m\times n}$, $B_0,\ B_1 \in \mathcal{R}^{m\times k}$, $C_0,\ C_1 \in \mathcal{R}^{l\times n}$, and $X_0,\ X_1 \in \mathcal{R}^{k\times l}$. In this chapter we determine maximal and minimal ranks of two real matrices X_0 and X_1 in solutions to the complex matrix equation in (25.1), and then present some consequences. To do so, we need the following result.

Lemma 25.1. *The complex matrix equation* (25.1) *is consistent if and only of the following real matrix equation*

$$\begin{bmatrix} B_0 & -B_1 \\ B_1 & B_0 \end{bmatrix} \begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix} \begin{bmatrix} C_0 & -C_1 \\ C_1 & C_0 \end{bmatrix} = \begin{bmatrix} A_0 & -A_1 \\ A_1 & A_0 \end{bmatrix}, \tag{25.2}$$

is consistent over the real number field $\mathcal{R}$. In that case the general solution of (25.1) *can be written as*

$$X = X_0 + iX_1 = \frac{1}{2}(Y_1 + Y_4) + \frac{i}{2}(Y_3 - Y_2), \tag{25.3}$$

where Y_1—Y_4 are the general solutions of (25.2) *over $\mathcal{R}$. Written in an explicit form, X_0 and X_1 in* (25.3) *are*

$$\begin{aligned} X_0 &= \frac{1}{2}P_1\phi^-(B)\phi(A)\phi^-(C)Q_1 + \frac{1}{2}P_2\phi^-(B)\phi(A)\phi^-(C)Q_2 \\ &\qquad + [\,P_1F_{\phi(B)},\ P_2F_{\phi(B)}\,]\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + [\,W_1,\ W_2\,]\begin{bmatrix} E_{\phi(C)}Q_1 \\ E_{\phi(C)}Q_2 \end{bmatrix}, \\ X_1 &= \frac{1}{2}P_2\phi^-(B)\phi(A)\phi^-(C)Q_1 - \frac{1}{2}P_1\phi^-(B)\phi(A)\phi^-(C)Q_2 \\ &\qquad + [\,P_2F_{\phi(B)},\ -P_1F_{\phi(B)}\,]\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + [\,W_1,\ W_2\,]\begin{bmatrix} -E_{\phi(C)}Q_2 \\ E_{\phi(C)}Q_1 \end{bmatrix}, \end{aligned}$$

where $\phi(M) = \phi(M_0 + iM_1) = \begin{bmatrix} M_0 & -M_1 \\ M_1 & M_0 \end{bmatrix}$, $P_1 = [\,I_k,\ 0\,]$, $P_2 = [\,0,\ I_k\,]$, $Q_1 = \begin{bmatrix} I_l \\ 0 \end{bmatrix}$, $Q_2 = \begin{bmatrix} 0 \\ I_l \end{bmatrix}$, V_1, V_2, W_1 *and* W_2 *are arbitrary over $\mathcal{R}$.*

Proof. It is well known that for any $M = M_0 + iM_1 \in \mathcal{C}^{m\times n}$, there is

$$\frac{1}{2}\begin{bmatrix} I_m & iI_m \\ -iI_m & -I_m \end{bmatrix}\begin{bmatrix} M_0 + iM_1 & 0 \\ 0 & M_0 - iM_1 \end{bmatrix}\begin{bmatrix} I_m & iI_m \\ -iI_m & -I_m \end{bmatrix} = \begin{bmatrix} M_0 & -M_1 \\ M_1 & M_0 \end{bmatrix} = \phi(M), \tag{25.4}$$

where $\phi(\cdot)$ satisfies the following operation properties

(i) $M = N \Leftrightarrow \phi(M) = \phi(N)$.
(ii) $\phi(M+N) = \phi(M)+\phi(N), \quad \phi(MN) = \phi(M)\phi(N), \quad \phi(kM) = k\phi(M), \ k \in \mathcal{R}$.
(iii) $\phi(M) = K_{2m}\phi(M)K_{2n}^{-1}$, where $K_{2t} = \begin{bmatrix} 0 & I_t \\ -I_t & 0 \end{bmatrix}, \quad t = m, \ n.$
(iv) $r[\phi(M)] = 2r(M)$.

Suppose now that (25.1) has a solution X over $\mathcal{C}$. Applying the above properties (i) and (ii) to it yields

$$\phi(B)\phi(X)\phi(C) = \phi(A), \tag{25.5}$$

which shows that $\phi(X)$ is a solution to (25.2). Conversely suppose that (25.2) has a solution $\widehat{Y} = \begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix} \in \mathcal{R}^{2k\times 2l}$, i.e., $\phi(B)\widehat{Y}\phi(C) = \phi(A)$. Then applying the above property (iii) to it yields

$$K_{2m}\phi(B)K_{2k}^{-1}\widehat{Y}K_{2l}\phi(C)K_{2n}^{-1} = K_{2m}\phi(A)K_{2n}^{-1},$$

consequently

$$\phi(B)(K_{2k}^{-1}\widehat{Y}K_{2l})\phi(C) = \phi(A),$$

which shows that $K_{2k}^{-1}\widehat{Y}K_{2l}$ is a solution of (25.5), too. Thus $\frac{1}{2}(\widehat{Y} + K_{2k}\widehat{Y}K_{2l})$ is a solution of (25.2), and this solution has the form

$$\frac{1}{2}(\widehat{Y} + \frac{1}{2}K_{2k}^{-1}\widehat{Y}K_{2l}) = \frac{1}{2}\begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} Y_4 & -Y_3 \\ -Y_2 & Y_1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} Y_1+Y_4 & -(Y_3-Y_2) \\ Y_3-Y_2 & Y_1+Y_4 \end{bmatrix}.$$

Let $\widehat{X} = \frac{1}{2}(Y_1+Y_4) + \frac{i}{2}(Y_3-Y_2)$. Then $\phi(\widehat{X}) = \frac{1}{2}(\widehat{Y} + K_{2k}^{-1}\widehat{Y}K_{2l})$ is a solution of (25.5). Thus by the above property (i), we know that $\widehat{X}$ is a solution of (25.1). The above derivation shows that the two equations (25.1) and (25.2) have the same consistency condition and their solutions satisfy the equality (25.3). Observe that Y_1 —Y_4 in (25.2) can be written as

$$Y_1 = P_1YQ_1, \quad Y_2 = P_1YQ_2, \quad Y_3 = P_2YQ_1, \quad Y_4 = P_2YQ_2,$$

where $Y = \begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix}$, and the general solution of (25.2) can be written as

$$Y = \phi^-(B)\phi(A)\phi^-(C) + 2F_{\phi(B)}[V_1, \ V_2] + 2\begin{bmatrix} W_1 \\ W_2 \end{bmatrix}E_{\phi(C)}.$$

Hence

$$\begin{aligned}
Y_1 &= P_1YQ_1 = P_1\phi^-(B)\phi(A)\phi^-(C)Q_1 + 2P_1F_{\phi(B)}V_1 + 2W_1E_{\phi(C)}Q_1,\\
Y_2 &= P_1YQ_2 = P_1\phi^-(B)\phi(A)\phi^-(C)Q_2 + 2P_1F_{\phi(B)}V_2 + 2W_1E_{\phi(C)}Q_2,\\
Y_3 &= P_2YQ_1 = P_2\phi^-(B)\phi(A)\phi^-(C)Q_1 + 2P_2F_{\phi(B)}V_1 + 2W_2E_{\phi(C)}Q_1,\\
Y_4 &= P_2YQ_2 = P_2\phi^-(B)\phi(A)\phi^-(C)Q_2 + 2P_2F_{\phi(B)}V_2 + 2W_2E_{\phi(C)}Q_2.
\end{aligned}$$

Putting them in (25.3) yields the general expressions of the two real matrices X_0 and X_1. □

Theorem 25.2. *Suppose the matrix equation* (25.1) *is consistent, and denote*

$$S_0 = \{ X_0 \in \mathcal{R}^{k\times l} \mid B(X_0+iX_1)C = A \}, \quad S_1 = \{ X_1 \in \mathcal{R}^{k\times l} \mid B(X_0+iX_1)C = A \}. \tag{25.6}$$

Then
(a) *The maximal and the minimal ranks of* X_0 *are given by*

$$\max_{X_0\in S_0} r(X_0) = \min\left\{ k, \ \ l, \ \ k+l+r\begin{bmatrix} A_0 & -A_1 & B_0 \\ A_1 & A_0 & B_1 \\ C_0 & -C_1 & 0 \end{bmatrix} - 2r(B) - 2r(C) \right\}, \tag{25.7}$$

$$\min_{X_0\in S_0} r(X_0) = r\begin{bmatrix} A_0 & -A_1 & B_0 \\ A_1 & A_0 & B_1 \\ C_0 & -C_1 & 0 \end{bmatrix} - r\begin{bmatrix} B_0 \\ B_1 \end{bmatrix} - r[C_0, \ C_1]. \tag{25.8}$$

(b) *The maximal and the minimal ranks of* X_1 *are given by*

$$\max_{X_1 \in S_1} r(X_1) = \min \left\{ k, \quad l, \quad k + l + r \begin{bmatrix} A_0 & -A_1 & B_0 \\ A_1 & A_0 & B_1 \\ C_0 & C_1 & 0 \end{bmatrix} - 2r(B) - 2r(C) \right\}, \tag{25.9}$$

$$\min_{X_1 \in S_1} r(X_1) = r \begin{bmatrix} A_0 & -A_1 & B_0 \\ A_1 & A_0 & B_1 \\ C_0 & C_1 & 0 \end{bmatrix} - r \begin{bmatrix} B_0 \\ B_1 \end{bmatrix} - r[\, C_0, \ C_1 \,]. \tag{25.10}$$

Proof. Applying (19.14) and (19.15) to X_0 in (25.3) yields

$$\max_{X_0 \in S_0} r(X_0) = \min\{\, k, \quad l, \quad r(M) \,\}, \tag{25.11}$$

$$\min_{X_0 \in S_0} r(X_0) = r(M) - r[\, P_1 F_{\phi(B)}, \ P_2 F_{\phi(B)} \,] - r \begin{bmatrix} E_{\phi(C)} Q_1 \\ E_{\phi(C)} Q_2 \end{bmatrix}, \tag{25.12}$$

where

$$M = \begin{bmatrix} \frac{1}{2} P_1 \phi^-(B)\phi(A)\phi^-(C) Q_1 + \frac{1}{2} P_2 \phi^-(B)\phi(A)\phi^-(C) Q_2 & P_1 F_{\phi(B)} & P_2 F_{\phi(B)} \\ E_{\phi(C)} Q_1 & 0 & 0 \\ E_{\phi(C)} Q_2 & 0 & 0 \end{bmatrix}.$$

Note that $\phi(B)\phi^-(B)\phi(C) = \phi(A)$ and $\phi(A)\phi^-(C)\phi(C) = \phi(A)$. By (1.3), (1.4) and (1.5), it is not difficult but tedious to find that

$$\begin{aligned} r(M) &\\ = \ r & \begin{bmatrix} \frac{1}{2} P_1 \phi^-(B)\phi(A)\phi^-(C) Q_1 + \frac{1}{2} P_2 \phi^-(B)\phi(A)\phi^-(C) Q_2 & P_1 & P_2 & 0 & 0 \\ Q_1 & 0 & 0 & \phi(C) & 0 \\ Q_2 & 0 & 0 & 0 & \phi(C) \\ 0 & \phi(B) & 0 & 0 & 0 \\ 0 & 0 & \phi(B) & 0 & 0 \end{bmatrix} \\ & -2r[\phi(B)] - 2r[\phi(C)] \\ = \ r & \begin{bmatrix} 0 & P_1 & P_2 & 0 & 0 \\ Q_1 & 0 & 0 & \phi(C) & 0 \\ Q_2 & 0 & 0 & 0 & \phi(C) \\ 0 & \phi(B) & 0 & 0 & 0 \\ 0 & 0 & \phi(B) & 0 & \phi(A) \end{bmatrix} - 2r[\phi(B)] - 2r[\phi(C)] \\ = \ r & \begin{bmatrix} 0 & C_0 & -C_1 \\ B_0 & A_0 & -A_1 \\ C_1 & A_1 & A_0 \end{bmatrix} - r[\phi(B)] - r[\phi(C)] + k + l \\ = \ r & \begin{bmatrix} A_0 & -A_1 & B_0 \\ A_1 & A_0 & B_1 \\ C_0 & -C_1 & 0 \end{bmatrix} - 2r(B) - 2r(C) + k + l, \end{aligned}$$

$$\begin{aligned} r[\, P_1 F_{\phi(B)}, \ P_2 F_{\phi(B)} \,] &= r \begin{bmatrix} P_1 & P_2 \\ \phi(B) & 0 \\ 0 & \phi(B) \end{bmatrix} - 2r[\phi(B)] \\ &= r \begin{bmatrix} -A_1 & 0 & -A_0 \\ A_0 & 0 & -A_1 \\ 0 & A_0 & -A_1 \\ 0 & A_1 & A_0 \end{bmatrix} - 2r[\phi(B)] + k \\ &= r \begin{bmatrix} B_0 \\ B_1 \end{bmatrix} - r[\phi(B)] + k = r \begin{bmatrix} B_0 \\ B_1 \end{bmatrix} - 2r(B) + k, \end{aligned}$$

$$r \begin{bmatrix} E_{\phi(C)} Q_1 \\ E_{\phi(C)} Q_2 \end{bmatrix} = r \begin{bmatrix} \phi(C) & 0 & Q_1 \\ 0 & \phi(C) & Q_2 \end{bmatrix} - 2r[\phi(C)]$$

$$= r\begin{bmatrix} C_1 & C_0 & 0 & 0 \\ 0 & 0 & C_0 & -C_1 \\ -C_0 & C_1 & C_1 & C_0 \end{bmatrix} - 2r[\phi(C)] + l$$
$$= r[\, C_0, \; C_1\,] - r[\phi(C)] + l = r[\, C_0, \; C_1\,] - 2r(C) + l.$$

Putting them in (25.10) and (25.12) yields (25.7) and (25.8). Similarly we can establish (25.9) and (25.10). □

Below is a direct consequence of Theorem 25.2.

Corollary 25.3. *Suppose the matrix equation* (25.1) *is consistent. Then*

(a) *Eq.* (25.1) *has a real solution* $X \in \mathcal{R}^{k\times l}$ *if and only if*

$$r\begin{bmatrix} A_0 & -A_1 & B_0 \\ A_1 & A_0 & B_1 \\ C_0 & C_1 & 0 \end{bmatrix} = r\begin{bmatrix} B_0 \\ B_1 \end{bmatrix} + r[\, C_0, \; C_1\,]. \tag{25.13}$$

(b) *All the solutions of* (25.1) *are real if and only if*

$$r\begin{bmatrix} A_0 & -A_1 & B_0 \\ A_1 & A_0 & B_1 \\ C_0 & C_1 & 0 \end{bmatrix} = 2r(B) + 2r(C) - k - l. \tag{25.14}$$

(c) *Eq.* (25.1) *has a pure imaginary solution* $X = iX_1$, *where* $X_1 \in \mathcal{R}^{k\times l}$, *if and only if*

$$r\begin{bmatrix} A_0 & -A_1 & B_0 \\ A_1 & A_0 & B_1 \\ C_0 & -C_1 & 0 \end{bmatrix} = r\begin{bmatrix} B_0 \\ B_1 \end{bmatrix} + r[\, C_0, \; C_1\,]. \tag{25.15}$$

(d) *All the solutions of* (25.1) *are pure imaginary if and only if*

$$r\begin{bmatrix} A_0 & -A_1 & B_0 \\ A_1 & A_0 & B_1 \\ C_0 & -C_1 & 0 \end{bmatrix} = 2r(B) + 2r(C) - k - l. \tag{25.16}$$

We next consider extreme ranks of $A_0 - B_0X_0C_0$ and $A_1 - B_1X_1C_1$ with respect to the real matrices X_0 and X_1 in solution of $(\,B_0 + iB_1\,)(\,X_0 + iX_1\,)(\,C_0 + iC_1\,) = A_0 + iA_1$, the corresponding results can be used in the next section to determine the relationships of A and C, B and D in generalized inverse $(\,A + iB\,)^- = C + iD$.

Theorem 25.4. *Suppose the matrix equation* (25.1) *is consistent, and* S_0 *is defined by* (25.6). *Then*

(a) *The maximal rank of* $A_0 - B_0X_0C_0$ *is*

$$\max_{X_0 \in S_0} r(\,A_0 - B_0X_0C_0\,)$$
$$= \min\left\{ r[\,A_0, \; B_0\,], \quad r\begin{bmatrix} A_0 \\ C_0 \end{bmatrix}, \quad r\begin{bmatrix} -A_0 & B_0 & 0 & 0 & 0 \\ C_0 & 0 & 0 & C_0 & -C_1 \\ 0 & 0 & 0 & C_1 & C_0 \\ 0 & B_0 & -B_1 & A_0 & -A_1 \\ 0 & B_1 & B_0 & A_1 & A_0 \end{bmatrix} - 2r(B) - 2r(C) \right\}.$$

(b) *The minimal rank of* $A_0 - B_0X_0C_0$ *is*

$$\min_{X_0 \in S_0} r(\,A_0 - B_0X_0C_0\,) = r[\,A_0, \; B_0\,] + r\begin{bmatrix} A_0 \\ C_0 \end{bmatrix} + r\begin{bmatrix} -A_0 & B_0 & 0 & 0 & 0 \\ C_0 & 0 & 0 & C_0 & -C_1 \\ 0 & 0 & 0 & C_1 & C_0 \\ 0 & B_0 & -B_1 & A_0 & -A_1 \\ 0 & B_1 & B_0 & A_1 & A_0 \end{bmatrix}$$
$$- r\begin{bmatrix} A_0 & B_0 & 0 \\ C_0 & 0 & 0 \\ 0 & B_0 & -B_1 \\ 0 & B_1 & B_0 \end{bmatrix} - r\begin{bmatrix} C_0 & A_0 & 0 & 0 \\ B_0 & 0 & B_0 & -B_1 \\ 0 & 0 & B_1 & B_0 \end{bmatrix}.$$

Proof. Putting the general expression X_0 in (25.3) in $A_0 - B_0X_0C_0$ yields

$$\begin{aligned} A_0 - B_0X_0C_0 &= A_0 - \frac{1}{2}B_0P_1\phi^-(B)\phi(A)\phi^-(C)Q_1C_0 - \frac{1}{2}B_0P_2\phi^-(B)\phi(A)\phi^-(C)Q_2B_0 \\ &\quad - B_0[\, P_1F_{\phi(B)},\ P_2F_{\phi(B)}\,]\begin{bmatrix} V_1 \\ V_2 \end{bmatrix}C_0 - B_0[\, W_1,\ W_2\,]\begin{bmatrix} E_{\phi(C)}Q_1 \\ E_{\phi(C)}Q_2 \end{bmatrix}C_0 \\ &= N - B_0GVC_0 - B_0WHC_0. \end{aligned}$$

Then according to (19.3) and (19.4), we get

$$\begin{aligned} \max_{X_0\in S_0} r(\, A_0 - B_0X_0C_0\,) &= \max_{V,\,W} r(\, N - B_0GVC_0 - B_0WHC_0\,) \\ &= \min\left\{ r[\, N,\ B_0\,],\ \ r\begin{bmatrix} N \\ C_0 \end{bmatrix},\ \ r\begin{bmatrix} N & B_0G \\ HC_0 & 0 \end{bmatrix} \right\}, \end{aligned}$$

$$\begin{aligned} &\min_{X_0\in S_0} r(\, A_0 - B_0X_0C_0\,) \\ &= \min_{V,\,W} r(\, N - B_0GVC_0 - B_0WHC_0\,) \\ &= r[\, N,\ B_0\,] + r\begin{bmatrix} N \\ C_0 \end{bmatrix} + r\begin{bmatrix} N & B_0G \\ HC_0 & 0 \end{bmatrix} - r\begin{bmatrix} N & B_0G \\ C_0 & 0 \end{bmatrix} - r\begin{bmatrix} N & B_0 \\ HC_0 & 0 \end{bmatrix}. \end{aligned}$$

Simplyfying the rank equalities by Lemma 1.1 may eventually results in the two equalities in parts (a) and (b). The processes, however, are quite tedious and are therefore omitted them here. □

Theorem 25.5. *Suppose the matrix equation* (25.1) *is consistent, and* S_1 *is defined by* (25.6). *Then*

(a) *The maximal rank of* $A_1 + B_1X_1C_1$ *is*

$$\begin{aligned} &\max_{X_1\in S_1} r(\, A_1 + B_1X_1C_1\,) \\ &= \min\left\{ r[\, A_1,\ B_1\,],\ \ r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix},\ \ r\begin{bmatrix} A_1 & B_1 & 0 & 0 & 0 \\ C_1 & 0 & 0 & C_1 & C_0 \\ 0 & 0 & 0 & -C_0 & C_1 \\ 0 & B_1 & B_0 & -A_1 & -A_0 \\ 0 & -B_0 & B_1 & A_0 & -A_1 \end{bmatrix} - 2r(B) - 2r(C) \right\}. \end{aligned}$$

(b) *The minimal rank of* $A_1 + B_1X_1C_1$ *is*

$$\begin{aligned} \min_{X_1\in S_1} r(\, A_1 + B_1X_1C_1\,) &= r[\, A_1,\ B_1\,] + r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix} + r\begin{bmatrix} A_1 & B_1 & 0 & 0 & 0 \\ C_1 & 0 & 0 & C_1 & C_0 \\ 0 & 0 & 0 & -C_0 & C_1 \\ 0 & B_1 & B_0 & -A_1 & -A_0 \\ 0 & -B_0 & B_1 & A_0 & -A_1 \end{bmatrix} \\ &\quad - r\begin{bmatrix} A_1 & B_1 & 0 \\ C_1 & 0 & 0 \\ 0 & B_1 & B_0 \\ 0 & -B_0 & B_1 \end{bmatrix} - r\begin{bmatrix} A_1 & B_1 & 0 & 0 \\ C_1 & 0 & C_1 & C_0 \\ 0 & 0 & -C_0 & C_1 \end{bmatrix}. \end{aligned}$$

Proof. Writing (25.1) in the following equivalent form

$$(\, A_1 - iA_0\,)(\, X_1 - iX_0\,)(\, B_1 - iB_0\,) = -C_1 + iC_0,$$

and then applying Theorem 2.4 to it yields (a) and (b). □

Corollary 25.6. *Suppose the matrix equation* (25.1) *is consistent, and the two sets* S_0 *and* S_1 *are defined by* (25.6).

(a) *If $B_0X_0C_0 = A_0$ is consistent over $\mathcal{R}$, then*

$$\min_{X_0 \in S_0} r(A_0 - B_0X_0C_0) = r\begin{bmatrix} -A_0 & B_0 & 0 & 0 & 0 \\ C_0 & 0 & 0 & C_0 & -C_1 \\ 0 & 0 & 0 & C_1 & C_0 \\ 0 & B_0 & -B_1 & A_0 & -A_1 \\ 0 & B_1 & B_0 & A_1 & A_0 \end{bmatrix} - r\begin{bmatrix} B_0 & 0 \\ 0 & B_1 \\ B_1 & B_0 \end{bmatrix} - r\begin{bmatrix} C_0 & 0 & C_1 \\ 0 & C_1 & C_0 \end{bmatrix}.$$

(b) *If $B_1X_1C_1 = -A_1$ is consistent over $\mathcal{R}$, then*

$$\min_{X_1 \in S_1} r(A_1 + B_1X_1C_1) = r\begin{bmatrix} A_1 & B_1 & 0 & 0 & 0 \\ C_1 & 0 & 0 & C_1 & C_0 \\ 0 & 0 & 0 & -C_0 & C_1 \\ 0 & B_1 & B_0 & -A_1 & -A_0 \\ 0 & -B_0 & B_1 & A_0 & -A_1 \end{bmatrix} - r\begin{bmatrix} B_0 & 0 \\ 0 & B_1 \\ B_1 & B_0 \end{bmatrix} - r\begin{bmatrix} C_0 & 0 & C_1 \\ 0 & C_1 & C_0 \end{bmatrix}.$$

(c) *The two linear matrix equations*

$$(B_0 + iB_1)(X_0 + iX_1)(C_0 + iC_1) = A_0 + iA_1 \quad \text{and} \quad B_0X_0C_0 = A_0$$

have a common solution for $X_0 \in \mathcal{R}^{p\times q}$ if and only if

$$r\begin{bmatrix} -A_0 & B_0 & 0 & 0 & 0 \\ C_0 & 0 & 0 & C_0 & -C_1 \\ 0 & 0 & 0 & C_1 & C_0 \\ 0 & B_0 & -B_1 & A_0 & -A_1 \\ 0 & B_1 & B_0 & A_1 & A_0 \end{bmatrix} = r\begin{bmatrix} B_0 & 0 \\ 0 & B_1 \\ B_1 & B_0 \end{bmatrix} + r\begin{bmatrix} C_0 & 0 & C_1 \\ 0 & C_1 & C_0 \end{bmatrix}.$$

(d) *The two linear matrix equations*

$$(B_0 + iB_1)(X_0 + iX_1)(C_0 + iC_1) = A_0 + iA_1 \quad \text{and} \quad B_1X_1C_1 = -A_1$$

have a common solution for $X_1 \in \mathcal{R}^{p\times q}$ if and only if

$$r\begin{bmatrix} A_1 & B_1 & 0 & 0 & 0 \\ C_1 & 0 & 0 & C_1 & C_0 \\ 0 & 0 & 0 & -C_0 & C_1 \\ 0 & B_1 & B_0 & -A_1 & -A_0 \\ 0 & -B_0 & B_1 & A_0 & -A_1 \end{bmatrix} = r\begin{bmatrix} B_0 & 0 \\ 0 & B_1 \\ B_1 & B_0 \end{bmatrix} + r\begin{bmatrix} C_0 & 0 & C_1 \\ 0 & C_1 & C_0 \end{bmatrix}.$$

Proof. Follows from simplification of Theorems 25.4 and 25.5 under the assumptions of the corollary. □

Applying Theorem 25.2 to the real matrices C and D in in the inner inverse $(A + iB)^- = C + iD$, we get the following.

Theorem 25.7. *Let $N = A + iB \in \mathcal{C}^{m\times n}$ be given, and denote*

$$T_1 = \{ C \in \mathcal{R}^{n\times m} \mid C + iD \in \{N^-\} \}, \quad T_2 = \{ D \in \mathcal{R}^{n\times m} \mid C + iD \in \{N^-\} \}. \tag{25.17}$$

(a) *The maximal and the minimal ranks of C in* (25.17) *are*

$$\max_{C \in T_1} r(C) = \min\left\{ m, \quad n, \quad m + n + r(A) - r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} \right\},$$

$$\min_{C \in T_1} r(C) = r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} - r\begin{bmatrix} A \\ B \end{bmatrix} - r[A, \ B] + r(A).$$

(b) *The maximal and the minimal ranks of D in* (25.17) *are*

$$\max_{D\in T_2} r(D) = \min\left\{ m, \quad n, \quad m+n+r(B) - r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} \right\},$$

$$\min_{D\in T_2} r(D) = r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} - r\begin{bmatrix} A \\ B \end{bmatrix} - r[\,A, \ B\,] + r(B).$$

(c) *If $A = 0$, then*

$$\max_{C\in T_1} r(C) = \min\{\, m, \quad n, \quad m+n-2r(B)\,\}.$$

(d) *If $B = 0$, then*

$$\max_{D\in T_2} r(D) = \min\{\, m, \quad n, \quad m+n-2r(A)\,\}.$$

(e) *If $R(A) \subseteq R(N)$ and $R(A^*) \subseteq R(N^*)$, then*

$$\min_{C\in T_1} r(C) = r(A), \qquad \min_{D\in T_2} r(D) = r(B).$$

Proof. Follows from replacing $A, \ B$ and C all by $N = A + iB$ in Theorem 25.2. □

Corollary 25.8. *Let $N = A + iB \in \mathcal{C}^{m\times n}$ be given, T_1 and T_2 be defined in* (25.17).

(a) *N has a real generalized inverse if and only if*

$$r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\,A, \ B\,] - r(B).$$

(b) *N has a pure imaginary generalized inverse if and only if*

$$r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} = r\begin{bmatrix} A \\ B \end{bmatrix} + r[\,A, \ B\,] - r(A).$$

Proof. Follows directly from Theorem 25.4(a) and (b). □

Theorem 25.9. *Let $N = A + iB \in \mathcal{C}^{m\times n}$ be given, T_1 and T_2 be defined by* (25.17).

(a) *The maximal and the minimal ranks of $A - ACA$ are*

$$\max_{C\in T_1} r(\,A - ACA\,) = \min\left\{ r(A), \quad r(A) + r\begin{bmatrix} 0 & B \\ B & A \end{bmatrix} - r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} \right\},$$

$$\min_{C\in T_1} r(\,A - ACA\,) = r(A) + r\begin{bmatrix} 0 & B \\ B & A \end{bmatrix} + r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & B & A \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ B & A \end{bmatrix}.$$

(b) *The maximal and the minimal ranks of $B + BDB$ are*

$$\max_{D\in T_2} r(\,B + BDB\,) = \min\left\{ r(B), \quad r(B) + r\begin{bmatrix} 0 & A \\ A & B \end{bmatrix} - r\begin{bmatrix} B & -A \\ A & B \end{bmatrix} \right\},$$

$$\min_{D\in T_2} r(\,B + BDB\,) = r(B) + r\begin{bmatrix} 0 & A \\ A & B \end{bmatrix} + r\begin{bmatrix} B & -A \\ A & B \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & B & A \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ B & A \end{bmatrix}.$$

Proof. Follows from replacing $A, \ B$ and C all by $N = A + iB$ in Theorems 25.4 and 25.5. □

Corollary 25.10. *Let $N = A + iB \in \mathcal{C}^{m\times n}$ be given, T_1 and T_2 be defined by* (25.17).

(a) *N has a generalized inverse with the form $N^- = A^- + iD$ if and only if*

$$r\begin{bmatrix} A & 0 & B \\ 0 & B & A \end{bmatrix} + r\begin{bmatrix} A & 0 \\ 0 & B \\ B & A \end{bmatrix} = r(A) + r\begin{bmatrix} 0 & B \\ B & A \end{bmatrix} + r\begin{bmatrix} A & -B \\ B & A \end{bmatrix}.$$

(b) $T_1 \subseteq \{A^-\}$, *i.e., all the generalized inverses of* N *have the form* $N^- = A^- + iD$ *if and only if*

$$r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} = r(A) + r\begin{bmatrix} 0 & B \\ B & A \end{bmatrix}.$$

(c) N *has a generalized inverse with the form* $N^- = C - iB^-$ *if and only if*

$$r\begin{bmatrix} A & 0 & B \\ 0 & B & A \end{bmatrix} + r\begin{bmatrix} A & 0 \\ 0 & B \\ B & A \end{bmatrix} = r(B) + r\begin{bmatrix} 0 & A \\ A & B \end{bmatrix} + r\begin{bmatrix} A & -B \\ B & A \end{bmatrix}.$$

(d) $T_2 \subseteq \{-B^-\}$, *i.e., all the generalized inverses of* N *have the form* $N^- = C - iB^-$ *if and only if*

$$r\begin{bmatrix} A & -B \\ B & A \end{bmatrix} = r(B) + r\begin{bmatrix} 0 & A \\ A & B \end{bmatrix}.$$

Proof. Follows directly from Theorem 25.9. $\square$

Chapter 26

Ranks and independence of solutions of the matrix equation $B_1XC_1 + B_2YC_2 = A$

We consider in the chapter possible ranks of solutions X and Y of the matrix equation
Suppose

$$B_1XC_1 + B_2YC_2 = A, \tag{26.1}$$

is a consistent linear matrix over an arbitrary field $\mathcal{F}$, where $B_1,\ C_1,\ B_2,\ C_2$ and A are $m \times p$, $q \times n$, $m \times s$, $t \times n$ and $m \times n$ matrices, respectively. In this chapter, We consider the maximal and the minimal ranks of solutions X and Y of (26.1), as well as independence of solutions X and Y of (26.1).

As one of basic linear matrix equations, (26.1) has been well examined in matrix theory and its applications (see, e.g., [5, 28, 69, 72, 110, 116, 131, 155]). Its solvability conditions and general solutions for X and Y are completely established by using ranks and generalized inverse of matrices. On the basis of those results and the rank formulas in the previous chapters, we now can give complete solutions to the above two problems.

The basic tools for investigating the above problems are the following several known results on ranks and generalized inverses of matrices.

Lemma 26.1. *Let $A \in \mathcal{F}^{m\times n}$, $B_1 \in \mathcal{F}^{m\times k_1}$, $B_2 \in \mathcal{F}^{m\times k_2}$, $C_1 \in \mathcal{F}^{l_1\times n}$ and $C_2 \in \mathcal{F}^{l_2\times n}$ be given. Then*

$$\max_{X,\,Y,\,Z} r(\, A - B_1XC_1 - B_2Y - ZC_2\,) = \min\left\{ m, \quad n, \quad r\begin{bmatrix} A & B_1 & B_2 \\ C_2 & 0 & 0 \end{bmatrix}, \quad r\begin{bmatrix} A & B_2 \\ C_1 & 0 \\ C_2 & 0 \end{bmatrix} \right\}, \tag{26.2}$$

$$\begin{aligned}\min_{X,\,Y,\,Z} r(\, A - B_1XC_1 - B_2Y - ZC_2\,) &= r\begin{bmatrix} A & B_1 & B_2 \\ C_2 & 0 & 0 \end{bmatrix} + r\begin{bmatrix} A & B_2 \\ C_1 & 0 \\ C_2 & 0 \end{bmatrix} \\ &\quad - r\begin{bmatrix} A & B_1 & B_2 \\ C_1 & 0 & 0 \\ C_2 & 0 & 0 \end{bmatrix} - r(B_2) - r(C_2). \end{aligned} \tag{26.3}$$

This lemma can be simply derived from the rank formulas in Theorem 18.4 and Corollary 19.5, and its proof is omitted here.

Concerning the general solution of (26.1), the following is well known.

Lemma 26.2. *Suppose the matrix equation is given by* (26.1). *Then*

(a)[131] *The general solution of the homogeneous equation $B_1XC_1 + B_2YC_2 = 0$ can factor as*

$$X = X_1X_2 + X_3, \qquad Y = Y_1Y_2 + Y_3,$$

where X_1—X_3 and Y_1—Y_3 are the general solutions of the following four simple homogeneous matrix equations

$$B_1X_1 = -B_2Y_1, \qquad X_2C_1 = Y_2C_2, \qquad B_1X_3C_1 = 0, \qquad B_2Y_3C_2 = 0, \tag{26.4}$$

Solving these four equations and putting their general solutions in X and Y yields

$$X = S_1F_GUE_HT_1 + F_{B_1}V_1 + V_2E_{C_1}, \qquad Y = S_2F_GUE_HT_2 + F_{B_2}W_1 + W_2E_{C_2},$$

where $S_1 = [\, I_p, \ 0 \,]$, $S_2 = [\, 0, \ I_s \,]$, $T_1 = \begin{bmatrix} I_q \\ 0 \end{bmatrix}$, $T_2 = \begin{bmatrix} 0 \\ I_t \end{bmatrix}$, $G = [\, B_1, \ B_2 \,]$ *and* $H = \begin{bmatrix} C_1 \\ -C_2 \end{bmatrix}$; *the matrices U, V_1, V_2, W_1 and W_2 are arbitrary.*

(b)[131] *Suppose the matrix equation* (26.1) *is consistent. Then its general solution can factor as*

$$X = X_0 + X_1X_2 + X_3, \qquad Y = Y_0 + Y_1Y_2 + Y_3,$$

where X_0 and Y_0 are a pair of particular solutions to (26.1), *X_1—X_3 and Y_1—Y_3 are the general solutions of the four simple matrix equations in* (26.4). *Written in an explicit form, the general solution of* (26.1) *is*

$$X = X_0 + S_1F_GUE_HT_1 + F_{B_1}V_1 + V_2E_{C_1}, \tag{26.5}$$
$$Y = Y_0 + S_2F_GUE_HT_2 + F_{B_2}W_1 + W_2E_{C_2}. \tag{26.6}$$

Various expressions of a pair of particular solutions of (26.1) can be found in [28], [69], [72], [110] and [155]. However we only use X_0 and Y_0 in form when determining possible ranks of solutions to (26.1), we do not intend to present their explicit expressions in (25.5) and (25.6).

For convenience of representation, we adopt the following notation

$$J_1 = \{\, X \in \mathcal{F}^{p\times q} \mid B_1XC_1 + B_2YC_2 = A \,\}, \quad J_2 = \{\, Y \in \mathcal{F}^{s\times t} \mid B_1XC_1 + B_2YC_2 = A \,\}. \tag{26.7}$$

The two expressions in (26.5) and (26.6) clearly show that the general solution X and Y of (26.1) are in fact two linear matrix expressions, each of which involves three independent variant matrices. In that case, apply Lemma 26.1 to obtain the following.

Theorem 26.3. *Suppose that the matrix equation* (26.1) *is consistent, and J_1 and J_2 are defined in* (26.7). *Then*

(a) *The maximal and the minimal ranks of solution X of* (26.1) *are*

$$\max_{X\in J_1} r(X) = \min\left\{p, \ q, \ p+q+r[\, A, \ B_2\,] - r[\, B_1, \ B_2\,] - r(C_1), \ p+q+r\begin{bmatrix} A \\ C_2 \end{bmatrix} - r\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} - r(B_1)\right\}, \tag{26.8}$$

$$\min_{X\in J_1} r(X) = r[\, A, \ B_2\,] + r\begin{bmatrix} A \\ C_2 \end{bmatrix} - r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix}. \tag{26.9}$$

(b) *The maximal and the minimal ranks of solution Y of* (26.1) *are*

$$\max_{Y\in J_2} r(Y) = \min\left\{s, \ t, \ s+t+r[\, A, \ B_1\,] - r[\, B_2, \ B_1\,] - r(C_2), \ s+t+r\begin{bmatrix} A \\ C_1 \end{bmatrix} - r\begin{bmatrix} C_2 \\ C_1 \end{bmatrix} - r(B_2)\right\}, \tag{26.10}$$

$$\min_{Y\in J_2} r(Y) = r[\, A, \ B_1\,] + r\begin{bmatrix} A \\ C_1 \end{bmatrix} - r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix}. \tag{26.11}$$

Proof. Applying (26.2) and (26.3) to (26.5) yields

$$\begin{aligned} \max_{X\in J_1} r(X) &= \max_{U,\, V_1,\, V_2} r(\, X_0 + S_1F_GUE_HT_1 + F_{B_1}V_1 + V_2E_{C_1} \,) \\ &= \min\left\{ p, \ q, \ r\begin{bmatrix} X_0 & F_{B_1} & S_1F_G \\ E_{C_1} & 0 & 0 \end{bmatrix}, \ r\begin{bmatrix} X_0 & F_{B_1} \\ E_{C_1} & 0 \\ E_HT_1 & 0 \end{bmatrix} \right\}, \end{aligned}$$

$$
\begin{aligned}
&\min_{X\in J_1} r(X)\\
&= \min_{U,\,V_1,\,V_2} r(\, X_0 + S_1F_GUE_HT_1 + F_{B_1}V_1 + V_2E_{C_1}\,)\\
&= r\begin{bmatrix} X_0 & F_{B_1} & S_1F_G\\ E_{C_1} & 0 & 0\end{bmatrix} + r\begin{bmatrix} X_0 & F_{B_1}\\ E_{C_1} & 0\\ E_HT_1 & 0\end{bmatrix} - r\begin{bmatrix} X_0 & F_{B_1} & S_1F_G\\ E_{C_1} & 0 & 0\\ E_HT_1 & 0 & 0\end{bmatrix} - r(F_{B_1}) - r(E_{C_1}).
\end{aligned}
$$

By Lemma 1.1 and $B_1X_0C_1 + B_2Y_0C_2 = A$, we find that $r(F_{B_1}) = p - r(B_1),\ r(E_{C_1}) = q - r(C_1)$, and

$$
\begin{aligned}
&r\begin{bmatrix} X_0 & F_{B_1} & S_1F_G\\ E_{C_1} & 0 & 0\end{bmatrix}\\
&= r\begin{bmatrix} X_0 & I_p & S_1 & 0\\ I_q & 0 & 0 & C_1\\ 0 & B_1 & 0 & 0\\ 0 & 0 & G & 0\end{bmatrix} - r(B_1) - r(C_1) - r(G)\\
&= r\begin{bmatrix} 0 & I_p & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & -B_1S_1 & B_1X_0C_1\\ 0 & 0 & G & 0\end{bmatrix} - r(B_1) - r(C_1) - r(G)\\
&= r\begin{bmatrix} -B_1 & 0 & B_1X_0C_1\\ B_1 & B_2 & 0\end{bmatrix} + p + q - r(B_1) - r(C_1) - r(G)\\
&= r[\, B_1,\ B_1X_0C_1\,] + p + q - r(C_1) - r(G) = r[\, B_2,\ A\,] + p + q - r(C_1) - r(G),
\end{aligned}
$$

$$
\begin{aligned}
&r\begin{bmatrix} X_0 & F_{B_1}\\ E_{C_1} & 0\\ E_HT_1 & 0\end{bmatrix}\\
&= r\begin{bmatrix} X_0 & I_p & 0 & 0\\ I_q & 0 & C_1 & 0\\ T_1 & 0 & 0 & H\\ 0 & B_1 & 0 & 0\end{bmatrix} - r(B_1) - r(C_1) - r(H)\\
&= r\begin{bmatrix} 0 & I_p & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & -T_1C & H\\ 0 & B_1X_0C & 0 & 0\end{bmatrix} - r(B_1) - r(C_1) - r(H)\\
&= r\begin{bmatrix} -C_1 & C_1\\ 0 & -C_2\\ B_1X_0C_1 & 0\end{bmatrix} + p + q - r(B_1) - r(C_1) - r(H)\\
&= r\begin{bmatrix} C_2\\ B_1X_0C_1\end{bmatrix} + p + q - r(B_1) - r(H) = r\begin{bmatrix} C_2\\ A\end{bmatrix} + p + q - r(B_1) - r(H),
\end{aligned}
$$

$$
\begin{aligned}
&r\begin{bmatrix} X_0 & F_{B_1} & S_1F_G\\ E_{C_1} & 0 & 0\\ E_HT_1 & 0 & 0\end{bmatrix}\\
&= r\begin{bmatrix} X_0 & I_p & S_1 & 0 & 0\\ I_q & 0 & 0 & C_1 & 0\\ T_1 & 0 & 0 & 0 & H\\ 0 & B_1 & 0 & 0 & 0\\ 0 & 0 & G & 0 & 0\end{bmatrix} - r(B_1) - r(C_1) - r(G) - r(H)\\
&= r\begin{bmatrix} 0 & I_p & 0 & 0 & 0\\ I_q & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & -T_1C & H\\ 0 & 0 & -B_1S_1 & B_1X_0C_1 & 0\\ 0 & 0 & G & 0 & 0\end{bmatrix} - r(B_1) - r(C_1) - r(G) - r(H)
\end{aligned}
$$

$$= r\begin{bmatrix} 0 & 0 & -C_1 & C_1 \\ 0 & 0 & 0 & -C_2 \\ -B_1 & 0 & B_1X_0C_1 & 0 \\ B_1 & B_2 & 0 & 0 \end{bmatrix} + p + q - r(B_1) - r(C_1) - r(G) - r(H)$$

$$= r\begin{bmatrix} 0 & 0 & -C_1 & 0 \\ 0 & 0 & 0 & -C_2 \\ -B_1 & 0 & 0 & 0 \\ 0 & B_2 & 0 & A \end{bmatrix} + p + q - r(B_1) - r(C_1) - r(G) - r(H)$$

$$= r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} + p + q - r(G) - r(H).$$

Thus we have the two formulas in Part (a). By the similar approach, we can establish Part (b). □

Furthermore, we can also find the maximal and the minimal ranks of B_1XC_1 and B_2YC_2 in (26.1) when it is consistent.

Theorem 26.4. *Suppose that the matrix equation* (26.1) *is consistent, and J_1 and J_2 are defined by* (26.7). *Then*

$$\max_{X \in J_1} r(B_1XC_1) = \min\left\{ r[\, A,\ B_2\,] - r[\, B_1,\ B_2\,] + r(B_1), \quad r\begin{bmatrix} A \\ C_2 \end{bmatrix} - r\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} + r(C_1) \right\}, \tag{26.12}$$

$$\min_{X \in J_1} r(B_1XC_1) = r[\, A,\ B_2\,] + r\begin{bmatrix} A \\ C_2 \end{bmatrix} - r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix}, \tag{26.13}$$

$$\max_{Y \in J_2} r(B_2YC_2) = \min\left\{ r[\, A,\ B_1\,] - r[\, B_2,\ B_1\,] + r(B_2), \quad r\begin{bmatrix} A \\ C_1 \end{bmatrix} - r\begin{bmatrix} C_2 \\ C_1 \end{bmatrix} + r(C_2) \right\}, \tag{26.14}$$

$$\min_{Y \in J_2} r(B_2YC_2) = r[\, A,\ B_1\,] + r\begin{bmatrix} A \\ C_1 \end{bmatrix} - r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix}. \tag{26.15}$$

Proof. Putting (26.4) to B_1XC_1 and then applying (18.5) and (18.6) , we find that

$$\max_{X \in J_1} r(B_1XC_1)$$
$$= \max_U r(\, B_1X_0C_1 + B_1S_1F_GUE_HT_1C_1\,) = \min\left\{ r[\, B_1X_0C_1,\ B_1S_1F_G\,], \quad r\begin{bmatrix} B_1X_0C_1 \\ E_HT_1C_1 \end{bmatrix} \right\},$$
$$\min_{X \in J_1} r(B_1XC_1)$$
$$= \min_U r(\, B_1X_0C_1 + B_1S_1F_GUE_HT_1C_1\,)$$
$$= r[\, B_1X_0C_1,\ B_1S_1F_G\,] + r\begin{bmatrix} B_1X_0C_1 \\ E_HT_1C_1 \end{bmatrix} - r\begin{bmatrix} B_1X_0C_1 & B_1S_1F_G \\ E_HT_1C_1 & 0 \end{bmatrix}.$$

Simplifying the ranks of the block matrices in them by Lemma 1.1 and $B_1X_0C_1 + B_2Y_0C_2 = A$, we get that

$$\begin{aligned} r[\, B_1X_0C_1,\ B_1S_1F_G\,] &= r\begin{bmatrix} B_1X_0C_1 & B_1S_1 \\ 0 & G \end{bmatrix} - r(G) \\ &= r\begin{bmatrix} B_1X_0C_1 & B_1 & 0 \\ 0 & B_1 & B_2 \end{bmatrix} - r(G) \\ &= r[\, B_1X_0C_1,\ B_2\,] + r(B_1) - r(G) = r[\, A,\ B_2\,] + r(B_1) - r(G), \end{aligned}$$

$$\begin{aligned} r\begin{bmatrix} B_1X_0C_1 \\ E_HT_1C_1 \end{bmatrix} &= r\begin{bmatrix} B_1X_0C_1 & 0 \\ T_1C_1 & H \end{bmatrix} - r(H) \\ &= r\begin{bmatrix} B_1X_0C_1 & 0 \\ C_1 & C_1 \\ 0 & -C_2 \end{bmatrix} - r(H) \\ &= r\begin{bmatrix} B_1X_0C_1 \\ C_2 \end{bmatrix} + r(C_1) - r(H) = r\begin{bmatrix} A \\ C_2 \end{bmatrix} + r(C_1) - r(H), \end{aligned}$$

$$
\begin{aligned}
r\begin{bmatrix} A_1X_0C_1 & A_1S_1F_G \\ E_HT_1C_1 & 0 \end{bmatrix} &= r\begin{bmatrix} B_1X_0C_1 & B_1S_1 & 0 \\ T_1C_1 & 0 & H \\ 0 & G & 0 \end{bmatrix} - r(G) - r(H) \\
&= r\begin{bmatrix} B_1X_0C_1 & B_1 & 0 & 0 \\ C_1 & 0 & 0 & C_1 \\ 0 & 0 & 0 & -C_2 \\ 0 & B_1 & B_2 & 0 \end{bmatrix} - r(G) - r(H) \\
&= r\begin{bmatrix} 0 & B_1 & 0 & 0 \\ C_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & C_2 \\ 0 & 0 & B_2 & B_1X_0C_1 \end{bmatrix} - r(G) - r(H) \\
&= r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} + r(B_1) + r(C_1) - r(G) - r(H).
\end{aligned}
$$

Therefore we have (26.12) and (26.13). Similarly we can show (26.14) and (26.15). □

Contrasting (26.9), (26.11), (26.13), (26.13) with (18.5), we find the following relations

$$
\begin{aligned}
&\min_{X \in J_1} r(X) = \min_{X \in J_1} r(B_1XC_1) = \min_{Y} r(\, A - B_2YC_2\,), \\
&\min_{Y \in J_2} r(Y) = \min_{Y \in J_2} r(B_2YC_2) = \min_{X} r(\, A - B_1XC_1\,).
\end{aligned}
$$

Theorem 26.5. *Suppose that the matrix equation* (26.1) *is consistent, and consider* J_1 *and* J_2 *in* (26.7) *as two independent matrix sets. Then*

$$
\max_{X \in J_1,\, Y \in J_2} r(\, A - B_1XC_1 - B_2YC_2\,) = \min\left\{ r(B_1) + r(B_2) - r[\, B_1,\ B_2\,], \quad r(C_1) + r(C_2) - r\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \right\}. \tag{26.16}
$$

In particular,

(a) *Solutions* X *and* Y *of* (26.1) *are independent, that is, for any* $X \in J_1$ *and* $Y \in J_2$ *the pair* X *and* Y *satisfy* (26.1)*, if and only if*

$$
R(B_1) \cap R(B_2) = \{0\}, \quad \textit{or} \quad R(C_1^T) \cap R(C_2^T) = \{0\}, \tag{26.17}
$$

where $R(\cdot)$ *denotes the column space of a matrix.*

(b) *Under* (26.17)*, the general solution of* (26.1) *can be written as the two independent forms*

$$
X = X_0 + S_1Q_GU_1P_HT_1 + F_{B_1}V_1 + V_2E_{C_1}, \qquad Y = Y_0 + S_2Q_GU_2P_HT_2 + F_{B_2}W_1 + W_2E_{C_2}, \tag{26.18}
$$

where X_0 *and* Y_0 *are a pair of special solutions of* (26.1)*,* U_1*,* U_2*,* V_1*,* V_2*,* W_1 *and* W_2 *are arbitrary.*

Proof. Writing (26.5) and (26.6) as two independent matrix expressions, that is, replacing U in (26.5) and (26.6) by U_1 and U_2 respectively, and then putting them in $A - B_1XC_1 - B_2YC_2$ yields

$$
\begin{aligned}
A - B_1XC_1 - B_2YC_2 &= A - B_1X_0C_1 - B_2Y_0C_2 - B_1S_1F_GU_1E_HT_1C_1 - B_2S_2F_GU_2E_HT_2C_2 \\
&= -B_1S_1F_GU_1E_HT_1C_1 - B_2S_2F_GU_2E_HT_2C_2 \\
&= -B_1S_1F_GU_1E_HT_1C_1 + B_1S_1F_GU_2E_HT_1C_1 \\
&= B_1S_1F_G(\, -U_1 + U_2\,)E_HT_1C_1,
\end{aligned}
$$

where U_1 and U_2 are arbitrary. Then by (18.5), it follows that

$$
\begin{aligned}
\max_{X \in J_1,\, Y \in J_2} r(\, A - B_1XC_1 - B_2YC_2\,) &= \max_{U_1,\, U_2} r[\, B_1S_1F_G(\, -U_1 + U_2\,)E_HT_1C_1\,] \\
&= \min\{\, r(B_1S_1F_G), \quad r(E_HT_1C_1)\,\},
\end{aligned}
$$

where

$$
r(B_1S_1F_G) = r\begin{bmatrix} B_1S_1 \\ G \end{bmatrix} - r(G) = r\begin{bmatrix} B_1 & 0 \\ B_1 & B_2 \end{bmatrix} - r(G) = r(B_1) + r(B_2) - r(G),
$$

$$r(E_H T_1 C_1) = r[\, T_1C_1,\ H\,] - r(H) = r\begin{bmatrix} C_1 & C_1 \\ 0 & -C_2 \end{bmatrix} - r(H) = r(C_1) + r(C_2) - r(H).$$

Therefore, we have (26.16). The result in (26.17) follows directly from (26.16) and the solutions in (26.18) follow from (26.5) and (26.7). □

Chapter 27

More on extreme ranks of $A - B_1X_1C_1 - B_2X_2C_2$ and related topics

In Chapter 19 we have presented extreme ranks of a matrix expression

$$p(X_1,\ X_2) = A - B_1X_1C_1 - B_2X_2C_2, \tag{27.1}$$

with respect to X_1 and X_2 under some restrictions on the given matrices in it. In this chapter we get rid of the restrictions to determine the maximal and the minimal ranks of $p(X_1,\ X_2)$ with respect to X_1 and X_2.

In the two papers [37] by Johnson and [154] by Woerdeman, maximal and minimal rank completions of partial banded block matrices were well examined. Two general methods for finding maximal and minimal ranks of partial banded block matrices were established in these two papers. According to the general methods, we can simply find the following two special results for a 3×3 partial banded block matrix.

Lemma 27.1[37][154]. *Let*

$$M = r\begin{bmatrix} A_{11} & A_{12} & X \\ A_{21} & A_{22} & A_{23} \\ Y & A_{32} & A_{33} \end{bmatrix}, \tag{27.2}$$

where $A_{ij} \in \mathcal{F}^{m_i \times n_j}$ $(1 \leq i,\ j \leq 3)$ *are given,* $X \in \mathcal{F}^{m_1 \times n_3}$ *and* $Y \in \mathcal{F}^{m_3 \times n_1}$ *are two variant matrices. Then*

$$\max_{X,\ Y} r(M) = \min\left\{ m_3 + n_3 + r\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad m_1 + n_1 + r\begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix}, \right.$$
$$\left. m_1 + m_3 + r[\, A_{21},\ A_{22},\ A_{23}\,], \quad n_1 + n_3 + r\begin{bmatrix} A_{12} \\ A_{22} \\ A_{32} \end{bmatrix} \right\}, \tag{27.3}$$

and

$$\min_{X,\ Y} r(M) = r[\, A_{21},\ A_{22},\ A_{23}\,] + r\begin{bmatrix} A_{12} \\ A_{22} \\ A_{32} \end{bmatrix} + \max\left\{ r\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} - r\begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} - r[\, A_{21},\ A_{22}\,], \right.$$
$$\left. r\begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix} - r\begin{bmatrix} A_{22} \\ A_{32} \end{bmatrix} - r[\, A_{22},\ A_{23}\,] \right\}. \tag{27.4}$$

Notice that the block matrix M in (27.2) and the matrix expression in (27.1) have two independent variant matrices, respectively. This fact motivates us to express the rank of (27.1) as the rank of a block matrix, and then apply (27.3) and (27.4) to determine extreme ranks of (27.1) with respect to X_1 and X_2.

It is easy to verify by block elementary operations of matrices that the rank of $p(X_1,\, X_2)$ in (27.1) satisfies the equality

$$r[\,p(X_1,\, X_2)\,] = r\begin{bmatrix} 0 & 0 & 0 & I_{p_2} & -X_2 \\ 0 & 0 & C_2 & 0 & I_{q_2} \\ 0 & B_1 & A & B_2 & 0 \\ I_{q_1} & 0 & C_1 & 0 & 0 \\ -X_1 & I_{p_1} & 0 & 0 & 0 \end{bmatrix} - p_1 - p_2 - q_1 - q_2. \tag{27.5}$$

Applying Lemma 1.1 to the block matrix in (27.5) and simplying, we obtain the main result of the chapter.

Theorem 27.2. *Let $p(X_1,\, X_2)$ be given by* (27.1). *Then*

$$\max_{X_1,\, X_2} r[\,p(X_1,\, X_2)\,] = \min\left\{ r[\,A,\; B_1,\; B_2\,], \quad r\begin{bmatrix} A \\ C_1 \\ C_2 \end{bmatrix}, \quad r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix}, \quad r\begin{bmatrix} A & B_2 \\ C_1 & 0 \end{bmatrix} \right\}, \tag{27.6}$$

and

$$\begin{aligned} &\min_{X_1,\, X_2} r[\,p(X_1,\, X_2)\,] \\ &= r\begin{bmatrix} A \\ C_1 \\ C_2 \end{bmatrix} + r[\,A,\; B_1,\; B_2\,] + \max\left\{ r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 & B_2 \\ C_2 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 \\ C_1 & 0 \\ C_2 & 0 \end{bmatrix}, \right. \\ &\qquad\qquad \left. r\begin{bmatrix} A & B_2 \\ C_1 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 & B_2 \\ C_1 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_2 \\ C_1 & 0 \\ C_2 & 0 \end{bmatrix} \right\}. \end{aligned} \tag{27.7}$$

The two rank equalities in (27.6) and (27.7) can help to reveal some fundamental properties of the matrix expression $p(X_1,\, X_2)$ in (27.1). For example, let

$$\max_{X_1,\, X_2} r[\,p(X_1,\, X_2)\,] = \min_{X_1,\, X_2} r[\,p(X_1,\, X_2)\,] = r(A),$$

one can immediately establish a necessary and sufficient condition for the rank of $p(X_1,\, X_2)$ to be invariant with respect to X_1 and X_2. Notice that two matrices M and N have the same column space if and only if $r[\,M,\; N\,] = r(M) = r(N)$. Thus the column space of $p(X_1,\, X_2)$ is invariant with respect to X_1 and X_2 if and only if

$$r[\,p(X_1,\, X_2),\; p(Y_1,\, Y_2)\,] = r[\,p(X_1,\, X_2)\,] = r[\,p(Y_1,\, Y_2)\,] = r(A)$$

holds for all $X_1,\; X_2,\; Y_1,\; Y_2$, where

$$\begin{aligned} [\,p(X_1,\, X_2),\; p(Y_1,\, Y_2)\,] &= [\,A - B_1X_1C_1 - B_2X_2C_2,\; A - B_1Y_1C_1 - B_2Y_2C_2\,] \\ &= [\,A,\; A\,] - B_1[\,X_1,\; Y_1\,]\begin{bmatrix} C_1 & 0 \\ 0 & C_1 \end{bmatrix} - B_2[\,X_2,\; Y_2\,]\begin{bmatrix} C_2 & 0 \\ 0 & C_2 \end{bmatrix}. \end{aligned}$$

Thus applying Theorem 27.1 to the equality, one can also establish a necessary and sufficient condition for the column space of $p(X_1,\, X_2)$ to be invariant with respect to X_1 and X_2. Moreover let (27.7) be zero, we can trivialy obtain a solvability condition for the matrix equation $B_1X_1C_1 + B_2X_2C_2 = A$, which has been established previously by Özgüler in [110].

Corollary 27.3. *There exist X_1 and X_2 such that $B_1X_1C_1 + B_2X_2C_2 = A$ if and only if*

$$r[\,A,\; B_1,\; B_2\,] = r[\,B_1,\; B_2\,], \qquad r\begin{bmatrix} A \\ C_1 \\ C_2 \end{bmatrix} = r\begin{bmatrix} C_1 \\ C_2 \end{bmatrix},$$

$$r\begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} = r(B_1) + r(C_2), \qquad r\begin{bmatrix} A & B_2 \\ C_1 & 0 \end{bmatrix} = r(B_2) + r(C_1).$$

Combining the two formulas (27.6), (26.7) and those in Chapter 19, we can also establish some more general results for linear matrix expressions with four two-sided independent variant matrices, which, in turn, will apply to determine extreme ranks for some more general matrix expressions.

Theorem 27.4. *Let*

$$p(X_1,\ X_2,\ X_3,\ X_4) = A - B_1X_1C_1 - B_2X_2C_2 - B_3X_3C_3 - B_4X_4C_4, \tag{27.8}$$

be a linear matrix expression with four two-sided terms over an arbitrary field $\mathcal{F}$, and suppose that the given matrices satisfying the conditions

$$R(B_i) \subseteq R(B_2), \quad and \quad R(C_j^T) \subseteq R(C_1^T), \qquad i = 1,\ 3,\ 4, \quad j = 2,\ 3,\ 4. \tag{27.9}$$

Then

$$\max_{X_i} r[\,p(X_1,\ X_2,\ X_3,\ X_4)\,] = \min\left\{ r[\,A,\ B_2\,],\ r\begin{bmatrix} A \\ C_1 \end{bmatrix},\ r\begin{bmatrix} A & B_1 \\ C_2 & 0 \\ C_3 & 0 \\ C_4 & 0 \end{bmatrix},\ r\begin{bmatrix} A & B_1 & B_3 & B_4 \\ C_2 & 0 & 0 & 0 \end{bmatrix}, \right.$$
$$\left. r\begin{bmatrix} A & B_1 & B_3 \\ C_2 & 0 & 0 \\ C_4 & 0 & 0 \end{bmatrix},\ r\begin{bmatrix} A & B_1 & B_4 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \end{bmatrix} \right\}, \tag{27.10}$$

and

$$\min_{X_i} r[\,p(X_1,\ X_2,\ X_3,\ X_4)\,]$$
$$= r\begin{bmatrix} A & B_1 \\ C_2 & 0 \\ C_3 & 0 \\ C_4 & 0 \end{bmatrix} + r\begin{bmatrix} A & B_1 & B_3 & B_4 \\ C_2 & 0 & 0 & 0 \end{bmatrix} + r\begin{bmatrix} A \\ C_1 \end{bmatrix} + r[\,A,\ B_2\,] - r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix}$$
$$+ \max\left\{ r\begin{bmatrix} A & B_1 & B_3 \\ C_2 & 0 & 0 \\ C_4 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 & B_3 & B_4 \\ C_2 & 0 & 0 & 0 \\ C_4 & 0 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 & B_3 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \\ C_4 & 0 & 0 \end{bmatrix}, \right.$$
$$\left. r\begin{bmatrix} A & B_1 & B_4 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 & B_3 & B_4 \\ C_2 & 0 & 0 & 0 \\ C_3 & 0 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 & B_4 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \\ C_4 & 0 & 0 \end{bmatrix} \right\}. \tag{27.11}$$

Proof. We only show (27.11). Under (27.9), we apply (19.4) to the two variant matrices X_1 and X_2 in (27.8) to yield

$$\min_{X_1,\, X_2} r[\,p(X_1,\ X_2,\ X_3,\ X_4)\,]$$
$$= r[\,A - B_3X_3C_3 - B_4X_4C_4,\ B_2\,] + r\begin{bmatrix} A - B_3X_3C_3 - B_4X_4C_4 \\ C_1 \end{bmatrix} + r\begin{bmatrix} A - B_3X_3C_3 - B_4X_4C_4 & B_1 \\ C_2 & 0 \end{bmatrix}$$
$$- r\begin{bmatrix} A - B_3X_3C_3 - B_4X_4C_4 & B_1 \\ C_1 & 0 \end{bmatrix} - r\begin{bmatrix} A - B_3X_3C_3 - B_4X_4C_4 & B_2 \\ C_2 & 0 \end{bmatrix}$$
$$= r[\,A,\ B_2\,] + r\begin{bmatrix} A \\ C_1 \end{bmatrix} - r\begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_2 \\ C_2 & 0 \end{bmatrix} + r\begin{bmatrix} A - B_3X_3C_3 - B_4X_4C_4 & B_1 \\ C_2 & 0 \end{bmatrix}. \tag{27.12}$$

Notice that

$$\begin{bmatrix} A - B_3X_3C_3 - B_4X_4C_4 & B_1 \\ C_2 & 0 \end{bmatrix} = \begin{bmatrix} A & B_1 \\ C_2 & 0 \end{bmatrix} - \begin{bmatrix} B_3 \\ 0 \end{bmatrix} X_3[\,C_3,\ 0\,] - \begin{bmatrix} B_4 \\ 0 \end{bmatrix} X_4[\,C_4,\ 0\,].$$

In that case, applying (27.7) to it and then putting the corresponding result in (27.12) yields (27.11). □

Corollary 27.5. *Let*

$$p(X_1,\ X_2,\ X_3,\ X_4) = A - B_1X_1 - X_2C_2 - B_3X_3C_3 - B_4X_4C_4 \tag{27.13}$$

be a linear matrix expression over an arbitrary field $\mathcal{F}$ with two one-sided terms and two two-sided terms. Then

$$\max_{\{X_i\}} r[\, p(X_1,\ X_2,\ X_3,\ X_4)\,] = \min\left\{ m, \quad n, \quad r\begin{bmatrix} A \\ C_1 \end{bmatrix}, \ r\begin{bmatrix} A & B_1 \\ C_2 & 0 \\ C_3 & 0 \\ C_4 & 0 \end{bmatrix}, \ r\begin{bmatrix} A & B_1 & B_3 & B_4 \\ C_2 & 0 & 0 & 0 \end{bmatrix}, \right.$$
$$\left. r\begin{bmatrix} A & B_1 & B_3 \\ C_2 & 0 & 0 \\ C_4 & 0 & 0 \end{bmatrix}, \ r\begin{bmatrix} A & B_1 & B_4 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \end{bmatrix} \right\}, \tag{27.14}$$

and

$$\min_{\{X_i\}} r[\, p(X_1,\ X_2,\ X_3,\ X_4)\,] = r\begin{bmatrix} A & B_1 \\ C_2 & 0 \\ C_3 & 0 \\ C_4 & 0 \end{bmatrix} + r\begin{bmatrix} A & B_1 & B_3 & B_4 \\ C_2 & 0 & 0 & 0 \end{bmatrix} - r(B_1) - r(C_2)$$
$$+ \max\left\{ r\begin{bmatrix} A & B_1 & B_3 \\ C_2 & 0 & 0 \\ C_4 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 & B_3 & B_4 \\ C_2 & 0 & 0 & 0 \\ C_4 & 0 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 & B_3 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \\ C_4 & 0 & 0 \end{bmatrix}, \right.$$
$$\left. r\begin{bmatrix} A & B_1 & B_4 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 & B_3 & B_4 \\ C_2 & 0 & 0 & 0 \\ C_3 & 0 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & B_1 & B_4 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \\ C_4 & 0 & 0 \end{bmatrix} \right\}. \tag{27.15}$$

In particular, the matrix equation

$$B_1X_1 + X_2C_2 + B_3X_3C_3 + B_4X_4C_4 = A \tag{27.16}$$

is consistent if and only if

$$r\begin{bmatrix} A & B_1 \\ C_2 & 0 \\ C_3 & 0 \\ C_4 & 0 \end{bmatrix} = r\begin{bmatrix} 0 & B_1 \\ C_2 & 0 \\ C_3 & 0 \\ C_4 & 0 \end{bmatrix}, \quad r\begin{bmatrix} A & B_1 & B_3 & B_4 \\ C_2 & 0 & 0 & 0 \end{bmatrix} = r\begin{bmatrix} 0 & B_1 & B_3 & B_4 \\ C_2 & 0 & 0 & 0 \end{bmatrix}, \tag{27.17}$$

$$r\begin{bmatrix} A & B_1 & B_3 \\ C_2 & 0 & 0 \\ C_4 & 0 & 0 \end{bmatrix} = r\begin{bmatrix} 0 & B_1 & B_3 \\ C_2 & 0 & 0 \\ C_4 & 0 & 0 \end{bmatrix}, \quad r\begin{bmatrix} A & B_1 & B_4 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \end{bmatrix} = r\begin{bmatrix} 0 & B_1 & B_4 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \end{bmatrix}. \tag{27.18}$$

Based on (27.10) and (27.11), we can determine extreme ranks of $A_1 - B_1XC_1$ subject to a pair of consistent matrix equations $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$. From them we can find lots of valuable results related to solvability and solutions of some matrix equations. We shall present them in the next chapter.

Some more general work than those in the chapter is to determine extreme ranks of a linear matrix expression $A - B_1X_1C_1 - B_2X_2C_2 - B_3X_3C_3$ with respect to $X_1,\ X_2$ and X_3, as well as $A - B_1X_1C_1 - \cdots - B_kX_kC_k,\ k > 3$ with respect to X_1—X_k without any restrictions to the given matrices in them. According to the method presented by Johnson in [75], the maximal rank of $A - B_1X_1C_1 - \cdots - B_kX_kC_k$ can completely be determined. Here we only list the case for $k = 3$ without its tedious proof.

Theorem 27.6. *Let*

$$p(\,X_1,\ X_2,\ X_3\,) = A - B_1X_1C_1 - B_2X_2C_2 - B_3X_3C_3$$

be a matrix expression over an arbitrary field $\mathcal{F}$. Then

$$\max_{\{X_i\}} r[\,p(X_1,\ X_2,\ X_3\,)\,] = \min\left\{ r\begin{bmatrix} A \\ C_1 \\ C_2 \\ C_3 \end{bmatrix},\ r\begin{bmatrix} A & B_1 \\ C_2 & 0 \\ C_3 & 0 \end{bmatrix},\ r\begin{bmatrix} A & B_2 \\ C_1 & 0 \\ C_3 & 0 \end{bmatrix},\ r\begin{bmatrix} A & B_3 \\ C_1 & 0 \\ C_2 & 0 \end{bmatrix},\right.$$

$$\left. r\begin{bmatrix} A & B_1 & B_2 \\ C_3 & 0 & 0 \end{bmatrix},\ r\begin{bmatrix} A & B_1 & B_3 \\ C_2 & 0 & 0 \end{bmatrix},\ r\begin{bmatrix} A & B_2 & B_3 \\ C_1 & 0 & 0 \end{bmatrix},\ r[\,A,\ B_1,\ B_2,\ B_3\,]\right\}. \tag{27.19}$$

¿From the eight block matrices in (27.19), the reader can easily infer the maximal rank of $A - B_1X_1C_1 - \cdots - B_kX_kC_k$, in which, the 2^k block matrices are much similar to those in (27.19).

As to the minimal rank of $A - B_1X_1C_1 - \cdots - B_kX_kC_k$ when $k \geq 3$, the process to find it becomes quite complicated. We do not find at present a general method to solve this challenging problem. However, as we have seen in Theorem 27.4 and Corollary 27.5, if the given matrices in a matrix expression satisfy some restrictions, then we can still find its minimal rank. Here we list two simple results.

Theorem 27.7. *Suppose $A_{ij}(1 \leq i,\ j \leq 3)$ are all nonsingular matrices of order m. Then*

$$\min_{X,\,Y,\,Z} r\begin{bmatrix} X & A_{12} & A_{13} \\ A_{21} & Y & A_{23} \\ A_{31} & A_{32} & Z \end{bmatrix} = \begin{cases} r(A)/2 & if\ r(A)\ is\ even \\ [\,r(A)+1\,]/2 & if\ r(A)\ is\ odd \end{cases}. \tag{27.20}$$

where

$$A = \begin{bmatrix} 0 & -A_{12} & -A_{13} \\ A_{21} & 0 & -A_{23} \\ A_{31} & A_{32} & 0 \end{bmatrix}. \tag{27.21}$$

Proof. According to (27.4), we first find that

$$\begin{aligned}
\min_{X,\,Z} r\begin{bmatrix} X & A_{12} & A_{13} \\ A_{21} & Y & A_{23} \\ A_{31} & A_{32} & Z \end{bmatrix} &= r[\,A_{21},\ Y,\ A_{23}\,] + r\begin{bmatrix} A_{12} \\ Y \\ A_{32} \end{bmatrix} \\
&\quad + \max\left\{ r\begin{bmatrix} A_{12} & A_{13} \\ Y & A_{23} \end{bmatrix} - r\begin{bmatrix} A_{12} \\ Y \end{bmatrix} - r[\,Y,\ A_{23}\,],\ \ r\begin{bmatrix} A_{21} & Y \\ A_{31} & A_{32} \end{bmatrix} - r\begin{bmatrix} Y \\ A_{32} \end{bmatrix} - r[\,A_{21},\ Y\,] \right\} \\
&= 2m + \max\left\{ r\begin{bmatrix} 0 & A_{13} \\ Y - A_{23}A_{13}^{-1}A_{12} & 0 \end{bmatrix} - 2m,\ \ r\begin{bmatrix} 0 & Y - A_{21}A_{31}^{-1}A_{32} \\ A_{31} & 0 \end{bmatrix} - 2m \right\} \\
&= m + \max\{\, r(\,Y - A_{23}A_{13}^{-1}A_{12}\,),\quad r(\,Y - A_{21}A_{31}^{-1}A_{32}\,)\,\} \\
&= m + \max\left\{\, r(\widehat{Y}),\quad r(\,\widehat{Y} - M\,)\,\right\},
\end{aligned}$$

where $\widehat{Y} = Y - A_{23}A_{13}^{-1}A_{12}$ and $M = A_{21}A_{31}^{-1}A_{32} - A_{23}A_{13}^{-1}A_{12}$. Thus

$$\min_{X,\,Y,\,Z} r\begin{bmatrix} X & A_{12} & A_{13} \\ A_{21} & Y & A_{23} \\ A_{31} & A_{32} & Z \end{bmatrix} = m + \min_{\widehat{Y}} \max\left\{\, r(\widehat{Y}),\quad r(\,M - \widehat{Y}\,)\,\right\}.$$

Notice that $r(\,M - \widehat{Y}\,) \geq r(M) - r(\widehat{Y})$ for all $\widehat{Y}$. We see that

$$\max\left\{\, r(\widehat{Y}),\quad r(\,M - \widehat{Y}\,)\,\right\} \geq \max\left\{\, r(\widehat{Y}),\quad r(M) - r(\widehat{Y})\,\right\},$$

and

$$\min_{\widehat{Y}} \max\left\{\, r(\widehat{Y}),\quad r(\,M - \widehat{Y}\,)\,\right\} \geq \min_{\widehat{Y}} \max\left\{\, r(\widehat{Y}),\quad r(M) - r(\widehat{Y})\,\right\}.$$

Since $\widehat{Y}$ is arbitrary, we easily get that

$$\min_{\widehat{Y}} \max\left\{\, r(\widehat{Y}),\quad r(M) - r(\widehat{Y})\,\right\} = \begin{cases} r(M)/2 & if\ r(M)\ is\ even \\ [\,r(M)+1\,]/2 & if\ r(M)\ is\ odd \end{cases}.$$

Consequently,

$$\min_{\widehat{Y}} \max\left\{ r(\widehat{Y}), \quad r(M - \widehat{Y}) \right\} \geq \left\{ \begin{array}{ll} r(M)/2 & if\ r(M)\ is\ even \\ [\, r(M) + 1\,]/2 & if\ r(M)\ is\ odd \end{array} \right. . \tag{27.22}$$

We next show that the lower bound in the right side pf (27.22) can be reached by the left hand side of (27.22) by choosing some $\widehat{Y}$. In fact, suppose M can factor as $M = P \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} Q$, where P and Q are nonsingular. If $k = r(M)$ is even, we take $\widehat{Y} = P \begin{bmatrix} I_{k/2} & 0 \\ 0 & 0 \end{bmatrix} Q$. In that case, $r(M - \widehat{Y}) = r(\widehat{Y}) = k/2$. If $k = r(M)$ is odd, then we take $\widehat{Y} = P \begin{bmatrix} I_{(k+1)/2} & 0 \\ 0 & 0 \end{bmatrix} Q$. In that case, $r(M - \widehat{Y}) = (k-1)/2$ and $r(\widehat{Y}) = (k+1)/2$. There two cases show that the left hand side of (27.22) can be reached by the right hand side of (27.22). Hence we have

$$\min_{\widehat{Y}} \max\left\{ r(\widehat{Y}), \quad r(M - \widehat{Y}) \right\} = \left\{ \begin{array}{ll} r(M)/2 & if\ r(M)\ is\ even \\ [\, r(M) + 1\,]/2 & if\ r(M)\ is\ odd \end{array} \right. .$$

Consequently,

$$\min_{X,\, Y,\, Z} r \begin{bmatrix} X & A_{12} & A_{13} \\ A_{21} & Y & A_{23} \\ A_{31} & A_{32} & Z \end{bmatrix} = m + \left\{ \begin{array}{ll} r(M)/2 & if\ r(M)\ is\ even \\ [\, r(M) + 1\,]/2 & if\ r(M)\ is\ odd \end{array} \right. . \tag{27.23}$$

On the other hand, it is easy to verify that

$$r \begin{bmatrix} 0 & -A_{12} & -A_{13} \\ A_{21} & 0 & -A_{23} \\ A_{31} & A_{32} & 0 \end{bmatrix} = r \begin{bmatrix} 0 & 0 & -A_{13} \\ 0 & A_{23}A_{13}^{-1}A_{12} - A_{21}A_{31}^{-1}A_{32} & 0 \\ A_{31} & 0 & 0 \end{bmatrix} = 2m + r(M).$$

Hence $r(M) = r(A) - 2m$. Putting it in (27.23) yields (27.20). □

Clearly the 3×3 block matrix in (27.20) is a special case of $A - B_1X_1C_1 - B_2X_2C_2 - B_3X_3C_3$. If A_{ij}'s are singular or are not square, the formula (27.20) is not valid. But we guess that its minimal rank can be expressed through the ranks of A and its submatrices.

Theorem 27.8. *Let*

$$p(X_1,\ \cdots,\ X_5\,) = A - B_1X_1C_1 - B_2X_2C_2 - B_3X_3C_3 - B_4X_4C_4 - B_5X_5C_5$$

be a matrix expression over an arbitrary field $\mathcal{F}$ with the given matrices satisfying the conditions

$$R[\, B_2,\ B_3\,] \subseteq R(B_1), \quad and \quad R[\, C_4^T,\ C_5^T\,] \subseteq R(C_1^T).$$

Then

$$\min_{X_i} r[\, p(X_1,\ \cdots,\ X_5\,)\,]$$
$$= \min_{X_4,\, X_5} r[\, A - B_4X_4C_4 - B_5X_5C_5,\ B_1\,] + \min_{X_2,\, X_3} r \begin{bmatrix} A - B_2X_2C_2 - B_3X_3C_3 \\ C_1 \end{bmatrix} - r \begin{bmatrix} A & B_1 \\ C_1 & 0 \end{bmatrix}. \tag{27.24}$$

Clearly the two minimal ranks in (27.24) can further be determined by (27.7). We leave it to the reader.

It should be mentioned that although failing to give the minimal rank of the matrix expression $A - B_1X_1C_1 - B_2X_2C_2 - B_3X_3C_3$ in general cases, we can still express consistency condition using rank equalities for the corresponding linear matrix equation

$$B_1X_1C_1 + B_2X_2C_2 + B_3X_3C_3 = A. \tag{27.25}$$

Here we only list the result, its proof will presented in chapter 28.

Theorem 27.9. *The matrix equation* (27.25) *is consistent if and only if the following nine rank equalities hold*

$$r\,[\,A,\ B_1,\ B_2,\ B_3\,] = r\,[\,B_1,\ B_2,\ B_3\,], \qquad r\begin{bmatrix} A \\ C_1 \\ C_2 \\ C_3 \end{bmatrix} = r\begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix},$$

$$r\begin{bmatrix} A & B_1 \\ C_2 & 0 \\ C_3 & 0 \end{bmatrix} = r\begin{bmatrix} C_2 \\ C_3 \end{bmatrix} + r(B_1), \qquad r\begin{bmatrix} A & B_2 \\ C_1 & 0 \\ C_3 & 0 \end{bmatrix} = r\begin{bmatrix} C_1 \\ C_3 \end{bmatrix} + r(B_2),$$

$$r\begin{bmatrix} A & B_3 \\ C_1 & 0 \\ C_2 & 0 \end{bmatrix} = r\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} + r(B_3), \qquad r\begin{bmatrix} A & B_1 & B_2 \\ C_3 & 0 & 0 \end{bmatrix} = r\,[\,B_1,\ B_2\,] + r(C_3),$$

$$r\begin{bmatrix} A & B_1 & B_3 \\ C_2 & 0 & 0 \end{bmatrix} = r\,[\,B_1,\ B_3\,] + r(C_2), \quad r\begin{bmatrix} A & B_2 & B_3 \\ C_1 & 0 & 0 \end{bmatrix} = r\,[\,B_2,\ B_3\,] + r(C_1),$$

$$r\begin{bmatrix} A & 0 & B_1 & 0 & B_3 \\ 0 & -A & 0 & B_2 & B_3 \\ C_2 & 0 & 0 & 0 & 0 \\ 0 & C_1 & 0 & 0 & 0 \\ C_3 & C_3 & 0 & 0 & 0 \end{bmatrix} = r\begin{bmatrix} C_2 & 0 \\ 0 & C_1 \\ C_3 & C_3 \end{bmatrix} + r\begin{bmatrix} B_1 & 0 & B_3 \\ 0 & B_2 & B_3 \end{bmatrix}.$$

Of course, those rank equalities can also equivalently be expressed by equivalence of matrices, column or row spaces of matrices, generalized inverses of matrices, and decompositions of matrices, and so on. The reader can easily list them according to Lemma 1.2.

Solvability and solutions of linear matrix equations have been one of principal topics in matrix theory and its applications. Based on the well-known Kronecker product of matrices, one can simply transform any kind of linear matrix equations to a standard form $Mx = b$, and then solve through it. Nearly all characteristics of the original equations, however, are lost in this kind of transformations. So one has been seeking various feasible methods to solve linear matrix equations without using the Kronecker product. As far as the author knows, Theorem 27.9 could be regarded as one of the most general conclusions on solvability of linear matrix equations up to now.

Just as what we did in Chapter 21, the two results (27.6) and (26.7) can be used to establish various types of rank equalities for generalized inverses of matrices. We next list some of them.

Theorem 27.10. *Let $A \in \mathcal{F}^{m\times k}$ and $B \in \mathcal{F}^{l\times m}$ be given. Then*

(a) *The maximal and the minimal ranks of $AA^- + B^-B$ with respect to A^- and B^- are*

$$\max_{A^-,\,B^-} r(\,AA^- + B^-B\,) = \min\{\ m,\ \ r(A) + r(B)\ \}, \tag{27.26}$$

$$\min_{A^-,\,B^-} r(\,AA^- + B^-B\,) = r(A) + r(B) - r(BA). \tag{27.27}$$

(b) *There are A^- and B^- such that $AA^- + B^-B$ is nonsingular if and only if $r(A) + r(B) \geq m$.*

(c) *The rank of $AA^- + B^-B$ is invariant with respect to the choice of A^- and B^- if and only if $BA = 0$ or $r(BA) = r(A) + r(B) - m$.*

(d) *The rank of $AA^- + B^-B$ is invariant with respect to the choice of A^- and B^- if and only if the rank of $AA^- - B^-B$ is invariant with respect to the choice of A^- and B^-.*

Proof. Note that

$$AA^- + B^-B = AA^\sim + B^\sim B + AV_1E_A + F_BV_2B.$$

This is a matrix expression with two independent variant matrices. Applying Theorem 27.2 to it and simplify we can get Part (a). The detailed is omitted here. Parts (b) and (c) are direct consequences of Part (a). Contracting Part (c) and Theorem 21.16(d) we get Part (d). □

Applying (27.27), (1.11) and (1.12), we can get the following

$$\min_{A^-,\,(I_m - A)^-} r[\,AA^- + (I_m - A)^-(I_m - A)\,] = r(A) + r(I_m - A) - r(A - A^2) = m, \tag{27.28}$$

$$\min_{(I_m+A)^-,\,(I_m-A)^-} r[\,(I_m+A)(I_m+A)^- + (I_m-A)^-(I_m-A)\,] = r(I_m+A) + r(I_m-A) - r(I_m-A^2) = m, \tag{27.29}$$

which imply that the matrices $AA^- + (I_m - A)^-(I_m - A)$ and $(I_m + A)(I_m + A)^- + (I_m - A)^-(I_m - A)$ are nonsingular for any A^-, $(I_m - A)^-$ and $(I_m + A)^-$.

By the similar approach, we can obtain the following.

Theorem 27.11. *Let $A \in \mathcal{F}^{m\times n}$ and $B \in \mathcal{F}^{m\times k}$ be given. Then*

(a) *The maximal and the minimal ranks of $AA^- + BB^-$ with respect to A^- and B^- are*

$$\max_{A^-,\,B^-} r(\,AA^- + BB^-\,) = r[\,A,\ B\,], \tag{27.30}$$

$$\min_{A^-,\,B^-} r(\,AA^- + BB^-\,) = \max\{\,r(A),\quad r(B)\,\}. \tag{27.31}$$

(b) *The maximal and the minimal ranks of $AA^- - BB^-$ with respect to A^- and B^- are*

$$\max_{A^-,\,B^-} r(\,AA^- - BB^-\,) = \min\{\,r[\,A,\ B\,],\quad r[\,A,\ B\,] + m - r(A) - r(B)\,\}, \tag{27.32}$$

$$\min_{A^-,\,B^-} r(\,AA^- - BB^-\,) = \max\{\,r[\,A,\ B\,] - r(A),\quad r[\,A,\ B\,] - r(B)\,\}. \tag{27.33}$$

(c) *There are A^- and B^- such that $AA^- = BB^-$ if and only if $R(A) = R(B)$.*

The rank equality in (27.30) can be extended to

$$\max_{A_1^-,\,\cdots,\,A_k^-} r(\,A_1A_1^- + \cdots + A_kA_k^-\,) = r[\,A_1,\ \cdots,\ A_k\,]. \tag{27.34}$$

Notice that $A_1A_1^- + \cdots + A_kA_k^-$ is in fact a matrix expression with k independent variant matrices. Hence we have no rank formula at present for determining the minimal rank of $A_1A_1^- + \cdots + A_kA_k^-$. Nevertheless, we can guess from (27.31) the following

$$\min_{A_1^-,\,\cdots,\,A_k^-} r(\,A_1A_1^- + \cdots + A_kA_k^-\,) = \max\{\,r(A_1),\ \cdots,\ r(A_k)\,\}. \tag{27.35}$$

Theorem 27.12. *Let $A \in \mathcal{F}^{n\times m}$ and $B \in \mathcal{F}^{k\times m}$ be given. Then*

$$\min_{A^-,\,B^-} r[\,A^-,\ B^-\,] = \min_{A^-,\,B^-} r[\,A^-A,\ B^-B\,] = \max\{\,r(A),\quad r(B)\,\}. \tag{27.36}$$

In general, we can guess from (27.36) the following

$$\min_{A_1^-,\,\cdots,\,A_k^-} r[\,A_1^-,\ \cdots,\ A_k^-\,] = \min_{A_1^-,\,\cdots,\,A_k^-} r[\,A_1^-A_1,\ \cdots,\ A_k^-A_k\,] = \max\{\,r(A_1),\ \cdots,\ r(A_k)\,\}. \tag{27.37}$$

Theorem 27.13. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ and $C \in \mathcal{F}^{l\times n}$ be given. Then*

$$\min_{B^-,\,C^-} r(\,A - BB^-A - AC^-C\,) \tag{27.38}$$

$$= \max\left\{ r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r(B) - r(C),\quad r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} + r(A) - r[\,A,\ B\,] - r\begin{bmatrix} A \\ C \end{bmatrix} \right\}. \tag{27.39}$$

In particular, there are B^- and C^- such that $BB^-A + AC^-C = A$, i.e., the matrix equation $BX + YC = A$ has a a solution with the form $X = B^-A$ and $Y = AC^-$, if and only if

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(B) + r(C) = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,] - r(A). \tag{27.40}$$

Moreover, one can also find extreme ranks of matrix expressions $A^kA^- + B^-B^k$, $A^kA^- \pm B^kB^-$, $A^-A^k \pm B^-B^k$, $A - BB^- \pm C^-C$, and so on. The reader can try them and establish some more general results.

Chapter 28

Extreme ranks of $A - B_1XC_1$ subject to $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$

This chapter considers extreme ranks of the matrix expression $A - B_1XC_1$ subject to a pair of consistent matrix equations $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ over an arbitrary filed $\mathcal{F}$. A direct motivation for this work comes from considering consistency of the triple matrix equations $B_1XC_1 = A_1,\ B_2XC_2 = A_2$ and $B_3XC_3 = A_3$. To do so, we need to know expression of general solution to the pair of matrix equations $B_2XC_2 = A_1,\ B_3XC_3 = A_3$.

Lemma 28.1. *Suppose that*

$$B_2XC_2 = A_2, \qquad B_3XC_3 = A_3 \tag{28.1}$$

is a pair of matrix equations over an arbitrary filed $\mathcal{F}$. Then

(a) *The general common solution of the pair of homogeneous matrix equations $B_2XC_2 = 0$ and $B_3XC_3 = 0$ can factor as*

$$X = X_1 + X_2 + X_3 + X_4, \tag{28.2}$$

where $X_1,\ X_2,\ X_3$ and X_4 are, respectively, the general solutions of the following four systems of homogeneous linear matrix equations

$$\left\{ \begin{array}{l} B_2X_1 = 0 \\ B_3X_1 = 0, \end{array} \right. \qquad \left\{ \begin{array}{l} X_2C_2 = 0 \\ X_2C_3 = 0, \end{array} \right. \qquad \left\{ \begin{array}{l} B_2X_3 = 0 \\ X_3C_3 = 0, \end{array} \right. \qquad \left\{ \begin{array}{l} X_4C_2 = 0 \\ B_3X_4 = 0. \end{array} \right. \tag{28.3}$$

Written in an explicit form, it is

$$X = F_BV_1 + V_2E_C + F_{B_2}V_3E_{C_3} + F_{B_3}V_4E_{C_2}, \tag{28.4}$$

where $B = \begin{bmatrix} B_2 \\ B_3 \end{bmatrix}$, $C = [\, C_2,\ C_3 \,]$, and V_1—V_4 are four arbitrary matrices.

(b) *Suppose that the pair of matrix equations $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ have a common solution. Then the general common solution can be written as*

$$X = X_0 + F_BV_1 + V_2E_C + F_{B_2}V_3E_{C_3} + F_{B_3}V_4E_{C_2}, \tag{28.5}$$

where X_0 is a particular common solution to $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$.

(c) *Suppose that the pair of matrix equations $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ have a common solution, and the given matrices satisfy*

$$R(B_2^T) \subseteq R(B_3^T), \qquad R(C_3) \subseteq R(C_2),$$

or equivalently

$$R(F_{B_3}) \subseteq R(F_{B_2}), \qquad R(E_{C_2}^T) \subseteq R(E_{C_3}^T).$$

Then the general common solution $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ can be written as

$$X = X_0 + F_{B_3}V_1 + V_2E_{C_2} + F_{B_2}V_3E_{C_3},$$

where X_0 is a particular common solution to $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$.

Proof. According to Lemma 18.1, the general solution of $B_2XC_2 = 0$ can be written as

$$X = F_{B_2}W_1 + W_2E_{C_2}, \tag{28.6}$$

where $W_1,\ W_2$ are arbitrary. Substituting it into $B_3XC_3 = 0$ yields

$$B_3XC_3 = B_3F_{B_2}W_1C_3 + B_3W_2E_{C_2}C_3 = 0. \tag{28.7}$$

Observe that $R(B_3F_{B_2}) \subseteq R(B_3)$ and $R[(E_{C_2}C_3)^T] \subseteq R(C_3^T)$. We can find by Lemma 26.2(a) that the general solutions for W_1 and W_2 of (28.7) can be written as

$$W_1 = UE_{C_2} + F_GV_1 + V_3E_{C_3},$$
$$W_2 = -F_{B_2}U + V_2E_H + F_{B_3}V_4,$$

where $H = B_3F_{B_2},\ G = E_{C_2}C_3$, and $U,\ V_1$—V_4 are arbitrary. Substituting both of them into (28.6) produces the general common solution of $B_2XC_2 = 0$ and $B_3XC_3 = 0$ as follows

$$X = F_{B_2}F_GV_1 + V_2E_HE_{C_2} + F_{B_2}V_3E_{C_3} + F_{B_3}V_4E_{C_2}. \tag{28.8}$$

It is easy to verify that the four terms in (28.8) are, in turn, the general common solutions of the four pairs of homogeneous equations in (28.3). Thus we have (28.2) and (28.4). The result in Part (b) is obvious from Part (a). □

Putting (28.5) in $A_1 - B_1XC_1$, we get

$$A_1 - B_1XC_1 = A_1 - B_1X_0C_1 - B_1F_BV_1C_1 - B_1V_2E_CC_1 - B_1F_{B_2}V_3E_{C_3}C_1 - B_1F_{B_3}V_4E_{C_2}C_1. \tag{28.9}$$

Thus the maximal and the minimal ranks of the matrix expression $A - B_1XC_1$ subject to $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ can be determined by the matrix expression (28.9). For convenience of representation, we write (28.9) as

$$A_1 - B_1XC_1 = A - G_1V_1H_1 - G_1V_2H_2 - G_3V_3H_3 - G_4V_4H_4, \tag{28.10}$$

where

$$A = A_1 - B_1X_0C_1, \quad G_1 = B_1F_B, \quad G_2 = B_1, \quad G_3 = B_1F_{B_2}, \quad G_4 = B_1F_{B_3}, \tag{28.11}$$
$$H_1 = C_1, \quad H_2 = E_CC_1, \quad H_3 = E_{C_3}C_1, \quad H_4 = E_{C_2}C_1. \tag{28.12}$$

Observe that (28.10) involves four independent variant matrices V_1—V_4. Moreover it is not difficult to derive that the above matrices satisfy the following conditions

$$R(G_1) \subseteq R(G_i) \subseteq R(G_2), \quad \text{and} \quad R(H_2^T) \subseteq R(H_i^T) \subseteq R(H_1^T), \qquad i = 3,\ 4, \tag{28.13}$$

Thus (28.10) can be regarded as a special case of the matrix expression in Theorem 27.4. In that case, applying the two formulas in (27.10) and (27.11) to (28.10), we get the main results of the chapter.

Theorem 28.2. *Suppose that the pair of matrix equations $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ have a common solution. Then the maximal rank of $A_1 - B_1XC_1$ subject to $A_2XB_2 = C_2$ and $A_3XB_3 = C_3$ is*

$$\max_{\substack{B_2XC_2 = A_2 \\ B_3XC_3 = A_3}} r(\, A_1 - B_1XC_1\,) = \min\left\{ r[\, A_1,\ B_1\,],\ \ r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix},\ \ s_1,\ \ s_2,\ \ s_3,\ \ s_4 \right\}, \tag{28.14}$$

where

$$s_1 = r\begin{bmatrix} A_1 & 0 & 0 & B_1 \\ 0 & -A_2 & 0 & B_2 \\ 0 & 0 & -A_3 & B_3 \\ C_1 & C_2 & 0 & 0 \\ C_1 & 0 & C_3 & 0 \end{bmatrix} - r\begin{bmatrix} B_2 \\ B_3 \end{bmatrix} - r(C_2) - r(C_3),$$

$$s_2 = r\begin{bmatrix} A_1 & 0 & 0 & B_1 & B_1 \\ 0 & -A_2 & 0 & B_2 & 0 \\ 0 & 0 & -A_3 & 0 & B_3 \\ C_1 & C_2 & C_3 & 0 & 0 \end{bmatrix} - r[\, C_2, \ C_4 \,] - r(B_2) - r(B_3),$$

$$s_3 = r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} - r(B_2) - r(C_2), \quad s_4 = r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_3 & B_3 \\ C_1 & C_3 & 0 \end{bmatrix} - r(B_3) - r(C_3).$$

Proof. Under (28.13), we first find by (27.10) that

$$\max_{\substack{B_2XC_2 = A_2 \\ B_3XC_3 = A_3}} r(\, A_1 - B_1XC_1 \,) = \max_{\{V_i\}} r(\, A - G_1V_1H_1 - G_1V_2H_2 - G_3V_3H_3 - G_4V_4H_4 \,)$$

$$= \min\left\{ r[\, A, \ G_2 \,], \ r\begin{bmatrix} A \\ H_1 \end{bmatrix}, \ r\begin{bmatrix} A & G_1 \\ H_3 & 0 \\ H_4 & 0 \end{bmatrix}, \ r\begin{bmatrix} A & G_3 & G_4 \\ H_2 & 0 & 0 \end{bmatrix}, \ r\begin{bmatrix} A & G_3 \\ H_4 & 0 \end{bmatrix}, \ r\begin{bmatrix} A & G_4 \\ H_3 & 0 \end{bmatrix} \right\}. \tag{28.15}$$

Simplifying the ranks of the block matrices in (28.15) by (1.2)—(1.4), as well as $B_2X_0C_2 = A_2$, $B_3X_0C_3 = A_3$, we have

$$r[\, A, \ G_2 \,] = r[\, A_1 - B_1X_0C_1, \ B_1 \,] = r[\, A_1, \ B_1 \,], \qquad r\begin{bmatrix} A \\ H_1 \end{bmatrix} = r\begin{bmatrix} A_1 - B_1X_0C_1 \\ C_1 \end{bmatrix} = r\begin{bmatrix} A_1 \\ C_1 \end{bmatrix},$$

$$r\begin{bmatrix} A & G_1 \\ H_3 & 0 \\ H_4 & 0 \end{bmatrix} = r\begin{bmatrix} A_1 - B_1X_0C_1 & B_1F_B \\ E_{C_3}C_1 & 0 \\ E_{C_2}C_1 & 0 \end{bmatrix}$$

$$= r\begin{bmatrix} A_1 - B_1X_0C_1 & B_1 & 0 & 0 \\ C_1 & 0 & C_3 & 0 \\ C_1 & 0 & 0 & C_2 \\ 0 & B_2 & 0 & 0 \\ 0 & B_3 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} B_2 \\ B_3 \end{bmatrix} - r(C_2) - r(C_3)$$

$$= r\begin{bmatrix} A_1 & B_1 & 0 & 0 \\ C_1 & 0 & C_3 & 0 \\ C_1 & 0 & 0 & C_2 \\ 0 & B_2 & 0 & -A_2 \\ 0 & B_3 & -A_3 & 0 \end{bmatrix} - r\begin{bmatrix} B_2 \\ B_3 \end{bmatrix} - r(C_2) - r(C_3)$$

$$= r\begin{bmatrix} A_1 & 0 & 0 & B_1 \\ 0 & -A_2 & 0 & B_2 \\ 0 & 0 & -A_3 & B_3 \\ C_1 & C_2 & 0 & 0 \\ C_1 & 0 & C_3 & 0 \end{bmatrix} - r\begin{bmatrix} B_2 \\ B_3 \end{bmatrix} - r(C_2) - r(C_3).$$

Similarly we can get

$$r\begin{bmatrix} A & G_3 & G_4 \\ H_2 & 0 & 0 \end{bmatrix} = r\begin{bmatrix} A_1 & 0 & 0 & B_1 & B_1 \\ 0 & -A_2 & 0 & B_2 & 0 \\ 0 & 0 & -A_3 & 0 & B_3 \\ C_1 & C_2 & C_3 & 0 & 0 \end{bmatrix} - r[\, C_2, \ C_4 \,] - r(B_2) - r(B_3),$$

$$r\begin{bmatrix} A & G_3 \\ H_4 & 0 \end{bmatrix} = r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} - r(B_2) - r(C_2),$$

$$r\begin{bmatrix} A & G_4 \\ H_3 & 0 \end{bmatrix} = r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_3 & B_3 \\ C_1 & C_3 & 0 \end{bmatrix} - r(B_3) - r(C_3).$$

Putting them in (28.15) yields (28.14). □

Theorem 28.3. *Suppose that the pair of matrix equations $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ have a common solution. Then the minimal rank of $A_1 - B_1XC_1$ subject to $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ is*

$$\begin{aligned}\min_{\substack{B_2XC_2 = A_2 \\ B_3XC_3 = A_3}} r(\, A_1 - B_1XC_1\,) &= r\begin{bmatrix} A_1 & 0 & 0 & B_1 \\ 0 & -A_2 & 0 & B_2 \\ 0 & 0 & -A_3 & B_3 \\ C_1 & C_2 & 0 & 0 \\ C_1 & 0 & C_3 & 0 \end{bmatrix} + r\begin{bmatrix} A_1 & 0 & 0 & B_1 & B_1 \\ 0 & -A_2 & 0 & B_2 & 0 \\ 0 & 0 & -A_3 & 0 & B_3 \\ C_1 & C_2 & C_3 & 0 & 0 \end{bmatrix} \\ &- r\begin{bmatrix} A & B_1 \\ C_1 & 0 \\ 0 & B_2 \\ 0 & B_3 \end{bmatrix} - r\begin{bmatrix} A & B_1 & 0 & 0 \\ C_1 & 0 & C_2 & C_3 \end{bmatrix} + \begin{bmatrix} A_1 \\ C_1 \end{bmatrix} + r[\, A_1, \ B_1\,] \\ &+ \max\left\{ r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} - r\begin{bmatrix} A_1 & 0 & B_1 & B_1 \\ 0 & -A_2 & B_2 & 0 \\ C_1 & C_2 & 0 & 0 \\ 0 & 0 & 0 & B_3 \end{bmatrix} - r\begin{bmatrix} A_1 & 0 & B_1 & 0 \\ 0 & -A_2 & B_2 & 0 \\ C_1 & C_2 & 0 & 0 \\ C_1 & 0 & 0 & C_3 \end{bmatrix}, \right. \\ &\left. r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_3 & B_3 \\ C_1 & C_3 & 0 \end{bmatrix} - r\begin{bmatrix} A_1 & 0 & B_1 & B_1 \\ 0 & -A_3 & B_3 & 0 \\ C_1 & C_3 & 0 & 0 \\ 0 & 0 & 0 & B_2 \end{bmatrix} - r\begin{bmatrix} A_1 & 0 & B_1 & 0 \\ 0 & -A_3 & B_3 & 0 \\ C_1 & C_3 & 0 & 0 \\ C_1 & 0 & 0 & C_2 \end{bmatrix} \right\}. \qquad (28.16)\end{aligned}$$

Proof. Under (28.13), applying (27.11) to (28.10) yields

$$\begin{aligned}&\min_{\substack{B_2XC_2 = A_2 \\ B_3XC_3 = A_3}} r(\, A_1 - B_1XC_1\,) \\ &= \min_{\{V_i\}} r(\, A - G_1V_1H_1 - G_1V_2H_2 - G_3V_3H_3 - G_4V_4H_4\,) \\ &= r[\, A, \ G_2\,] + r\begin{bmatrix} A \\ H_1 \end{bmatrix} + r\begin{bmatrix} A & G_1 \\ H_3 & 0 \\ H_4 & 0 \end{bmatrix} + r\begin{bmatrix} A & G_3 & G_4 \\ H_2 & 0 & 0 \end{bmatrix} + r[\, A, \ G_2\,] + r\begin{bmatrix} A \\ H_1 \end{bmatrix} - r\begin{bmatrix} A & G_1 \\ H_1 & 0 \end{bmatrix} \\ &\quad - r\begin{bmatrix} A & G_2 \\ H_2 & 0 \end{bmatrix} + \max\left\{ r\begin{bmatrix} A & G_3 \\ H_4 & 0 \end{bmatrix} - r\begin{bmatrix} A & G_3 & G_4 \\ H_3 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & G_3 \\ H_3 & 0 \\ H_4 & 0 \end{bmatrix}, \right. \\ &\quad \left. r\begin{bmatrix} A & G_4 \\ H_3 & 0 \end{bmatrix} - r\begin{bmatrix} A & G_3 & G_4 \\ H_3 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & G_4 \\ H_3 & 0 \\ H_4 & 0 \end{bmatrix} \right\}. \qquad (28.17)\end{aligned}$$

Simplifying the ranks of block matrices in (28.17) by (1.2)—(1.4), as well as $B_2X_0C_2 = A_2$, $B_3X_0C_3 = A_3$, we can eventually get the rank formula (28.16). But we omit here the tedious steps. □

Corollary 28.4. *Suppose that the three matrix equations $B_1XC_1 = A_1$, $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ are consistent, respectively. Also suppose that any pair of the three matrix equations has a common solution. Then*

$$\min_{\substack{B_2XC_2 = A_2 \\ B_3XC_3 = A_3}} r(\, A_1 - B_1XC_1\,) = r\begin{bmatrix} A_1 & 0 & 0 & B_1 \\ 0 & -A_2 & 0 & B_2 \\ 0 & 0 & -A_3 & B_3 \\ C_1 & C_2 & 0 & 0 \\ C_1 & 0 & C_3 & 0 \end{bmatrix} + r\begin{bmatrix} A_1 & 0 & 0 & B_1 & B_1 \\ 0 & -A_2 & 0 & B_2 & 0 \\ 0 & 0 & -A_3 & 0 & B_3 \\ C_1 & C_2 & C_3 & 0 & 0 \end{bmatrix}$$

$$- r\begin{bmatrix} B_1 & B_1 \\ B_2 & 0 \\ 0 & B_3 \end{bmatrix} - r\begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} - r\begin{bmatrix} C_1 & C_2 & 0 \\ C_1 & 0 & C_3 \end{bmatrix} - r[\, C_1, \ C_2, \ C_3 \,]. \tag{28.18}$$

Proof. Under the assumption of the corollary, we know by Corollary 20.3 that the given matrices in the three equations satisfy the conditions

$$R(A_i) \subseteq R(B_i), \quad R(A_i^T) \subseteq R(C_i^T), \quad i = 1, \ 2, \ 3, .$$

$$r\begin{bmatrix} A_i & 0 & B_i \\ 0 & -A_j & B_j \\ C_i & C_j & 0 \end{bmatrix} = r\begin{bmatrix} B_i \\ B_j \end{bmatrix} + r[\, C_i, \ C_j \,]. \quad i = 1, \ 2, \ 3.$$

In that case, the formula (28.16) reduces to (28.18). $\square$

Based on the formula (28.18), one can easily verify that under the assumption of Corollary 28.4, the following identity holds

$$\min_{\substack{B_2XC_2 = A_2 \\ B_3XC_3 = A_3}} r(\, A_1 - B_1XC_1 \,) = \min_{\substack{B_1XC_1 = A_1 \\ B_3XC_3 = A_3}} r(\, A_2 - B_2XC_2 \,) = \min_{\substack{B_1XC_1 = A_1 \\ B_2XC_2 = A_2}} r(\, A_3 - B_3XC_3 \,).$$

One of the most important consequences of (28.18) is concerning the consistency of a triple matrix equations.

Corollary 28.5. *The triple linear matrix equations $B_1XC_1 = A_1$, $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ have a common solution if and only if any pair of the three equations has a common solution, meanwhile the given matrices satisfy the two rank equalities*

$$r\begin{bmatrix} A_1 & 0 & 0 & B_1 & B_1 \\ 0 & -A_2 & 0 & B_2 & 0 \\ 0 & 0 & -A_3 & 0 & B_3 \\ C_1 & C_2 & C_3 & 0 & 0 \end{bmatrix} = r\begin{bmatrix} B_1 & B_1 \\ B_2 & 0 \\ 0 & B_3 \end{bmatrix} + r[\, C_1, \ C_2, \ C_3 \,], \tag{28.19}$$

$$r\begin{bmatrix} A_1 & 0 & 0 & B_1 \\ 0 & -A_2 & 0 & B_2 \\ 0 & 0 & -A_3 & B_3 \\ C_1 & C_2 & 0 & 0 \\ C_1 & 0 & C_3 & 0 \end{bmatrix} = r\begin{bmatrix} C_1 & C_2 & 0 \\ C_1 & 0 & C_3 \end{bmatrix} + r\begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}. \tag{28.20}$$

This result can also be alternatively stated as follows.

Corollary 28.6. *The triple matrix equations $B_1XC_1 = A_1$, $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ have a common solution if and only if the following eight independent simple matrix equations are all solvable*

$$B_1X_1C_1 = A_1, \quad B_2X_2C_2 = A_2, \quad B_3X_3C_3 = A_3,$$

$$\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} X_4 + Y_4[\, C_1, \ C_2 \,] = \begin{bmatrix} A_1 & 0 \\ 0 & -A_2 \end{bmatrix}, \quad \begin{bmatrix} B_1 \\ B_3 \end{bmatrix} X_5 + Y_5[\, C_1, \ C_3 \,] = \begin{bmatrix} A_1 & 0 \\ 0 & -A_3 \end{bmatrix},$$

$$\begin{bmatrix} B_2 \\ B_3 \end{bmatrix} X_6 + Y_6[\, C_2, \ C_3 \,] = \begin{bmatrix} A_2 & 0 \\ 0 & -A_3 \end{bmatrix},$$

$$\begin{bmatrix} B_1 & B_1 \\ B_2 & 0 \\ 0 & B_3 \end{bmatrix} X_7 + Y_7[\, C_1, \ C_2, \ C_3 \,] = \begin{bmatrix} A_1 & 0 & 0 \\ 0 & -A_2 & 0 \\ 0 & 0 & -A_3 \end{bmatrix},$$

$$\begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} X_8 + Y_8 \begin{bmatrix} C_1 & C_2 & 0 \\ C_1 & 0 & C_3 \end{bmatrix} = \begin{bmatrix} A_1 & 0 & 0 \\ 0 & -A_2 & 0 \\ 0 & 0 & -A_3 \end{bmatrix}.$$

Of course, one can also equivalently write the consistency condition for $B_1XC_1 = A_1$, $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ in Theorems 28.5 and 28.6 in term of equivalence of matrices, column or row spaces of matrices, generalized inverses of matrices, and so on.

As a simple consequence of Theorems 28.2 and 28.3 we can also get the maximal and the minimal ranks of common solutions to a pair of linear matrix equations. This problem was examined by Mitra [103].

Corollary 28.7. *Suppose that the pair of matrix equations $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ have a common solution, where X is a $p \times q$ matrix. Then the maximal rank of common solutions to the pair of equations is*

$$\max_{\substack{B_2XC_2 = A_2 \\ B_3XC_3 = A_3}} r(X) = \min\{ p, \quad q, \quad s_1, \quad s_2, \quad s_3, \quad s_4 \}, \tag{28.21}$$

where

$$\begin{aligned} s_1 &= r(A_2) - r(B_2) - r(C_2) + p + q, \\ s_2 &= r(A_3) - r(B_3) - r(C_3) + p + q, \\ s_3 &= r\begin{bmatrix} A_2 & 0 \\ 0 & A_3 \\ C_2 & C_3 \end{bmatrix} - r[\, C_2, \ C_3\,] - r(C_2) - r(C_3) + p + q, \\ s_4 &= r\begin{bmatrix} A_2 & 0 & B_2 \\ 0 & A_3 & B_3 \end{bmatrix} - r\begin{bmatrix} B_2 \\ B_3 \end{bmatrix} - r(B_2) - r(B_3) + p + q. \end{aligned}$$

The minimal rank of common solutions to the pair of equations is

$$\min_{\substack{B_2XC_2 = A_2 \\ B_3XC_3 = A_3}} r(X) = r\begin{bmatrix} A_2 & 0 \\ 0 & A_3 \\ C_2 & C_3 \end{bmatrix} + r\begin{bmatrix} A_2 & 0 & B_2 \\ 0 & A_3 & B_3 \end{bmatrix}$$

$$+ \max\left\{ r(A_2) - r\begin{bmatrix} A_2 & B_2 \\ 0 & B_3 \end{bmatrix} - r\begin{bmatrix} A_2 & 0 \\ C_2 & C_3 \end{bmatrix}, \ \ r(A_3) - r\begin{bmatrix} B_2 & 0 \\ B_3 & A_3 \end{bmatrix} - r\begin{bmatrix} C_2 & C_3 \\ 0 & A_3 \end{bmatrix} \right\}. \tag{28.22}$$

Corollary 28.8. *Let $A, \ B, \ C \in \mathcal{F}^{m \times n}$ be given. Then $A, \ B$ and C have a common inner inverse if and only if*

$$\begin{aligned} r(\, A - B\,) &= r\begin{bmatrix} A \\ B \end{bmatrix} + r[\, A, \ B\,] - r(A) - r(B), \\ r(\, A - C\,) &= r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A, \ C\,] - r(A) - r(C), \\ r(\, B - C\,) &= r\begin{bmatrix} B \\ C \end{bmatrix} + r[\, B, \ C\,] - r(B) - r(C), \\ r[\, A - B, \ A - C] &= r\begin{bmatrix} A & A \\ B & 0 \\ 0 & C \end{bmatrix} + r[\, A, \ B, \ C\,] - r(A) - r(B) - r(C), \\ r\begin{bmatrix} A - B \\ A - C \end{bmatrix} &= r\begin{bmatrix} A \\ B \\ C \end{bmatrix} + r\begin{bmatrix} A & B & 0 \\ A & 0 & C \end{bmatrix} - r(A) - r(B) - r(C). \end{aligned}$$

In particular, if

$$r\begin{bmatrix} A \\ B \\ C \end{bmatrix} = r[\, A, \ B, \ C\,] = r(A) + r(B) + r(C),$$

then $A, \ B$ and C have a common inner inverse.

Proof. Consider the three matrix equations $AXA = A$, $BXB = B$ and $CXC = C$. Then the result in the corollary follows directly from Theorems 21.10(a) and 28.5. $\square$

When the matrices A, B and C are all idempotent, they have identity matrix as their common inner inverse. Thus the five rank equalities in Corollary 28.7 are all satisfied, the first three occurred in Theorem 3.1, the fourth and the fifth are two new rank equalities for idempotent matrices.

Another work related to a triple matrix equations $B_1XC_1 = A_1$, $B_2XC_2 = A_2$ and $B_3XC_3 = A_3$ is to determine

$$\max_{\substack{B_1XC_1 = A_1 \\ B_2YC_2 = A_2 \\ B_3YC_3 = A_3}} r(\, X - Y\,), \qquad \min_{\substack{B_1XC_1 = A_1 \\ B_2YC_2 = A_2 \\ B_3YC_3 = A_3}} r(\, X - Y\,). \tag{28.23}$$

Based on Lemma 28.1 and Corollary 27.5, one can routinely find the two ranks in (28.23). From them one can also establish Corollaries 28.5 and 28.6. We leave this work to the reader.

A more general work than those for a triple matrix equations is to consider common solution to a quadruple of matrix equations

$$B_1XC_1 = A_1, \quad B_2XC_2 = A_2, \quad B_3XC_3 = A_3, \quad B_4XC_4 = A_4. \tag{28.24}$$

Clearly, the quadruple matrix equations have a common solution if and only if

$$\min_{\substack{B_1XC_1 = A_1 \\ B_2XC_2 = A_2 \\ B_3YC_3 = A_3 \\ B_4YC_4 = A_4}} r(\, X - Y\,) = 0. \tag{28.25}$$

If the two pairs $B_1XC_1 = A_1$, $B_2XC_2 = A_2$ and $B_3YC_3 = A_3$, $B_4YC_4 = A_4$ are consistent respectively, then the deference $X - Y$ in (28.25), according to Lemma 28.1(b), is a linear matrix expression with eight independent variant matrices, four of them are one-sided and other four are two-sided. Unfortunately we can not find in general the minimal rank of such a matrix expression. However if the quadruple matrix equations satisfy some restrictions, and the expressions for X and Y are reduced to some simple forms, then we can find (28.25). One such a case is when (28.24) satisfy the conditions

$$R(B_1^T) \subseteq R(B_2^T), \qquad R(C_2) \subseteq R(C_1), \qquad R(B_3^T) \subseteq R(B_4^T), \qquad R(C_4) \subseteq R(C_3). \tag{28.26}$$

or equivalently

$$R(F_{B_2}) \subseteq R(F_{B_1}), \qquad R(E_{C_1}^T) \subseteq R(E_{C_2}^T), \qquad R(F_{B_4}) \subseteq R(F_{B_3}), \qquad R(E_{C_3}^T) \subseteq R(E_{C_4}^T). \tag{28.27}$$

In that case, the general common solution to $B_1XC_1 = A_1$ and $B_2XC_2 = A_2$, according to Lemma 28.1(c), is

$$X = X_0 + F_{B_2}V_1 + V_2E_{C_1} + F_{B_1}V_3E_{C_2},$$

where X_0 is a particular common solution to the pair $B_1XC_1 = A_1$ and $B_2XC_2 = A_2$, V_1—V_3 are arbitrary; the general common solution to $B_3YC_3 = A_3$ and $B_4YC_4 = A_4$ is is

$$Y = Y_0 - F_{B_4}W_1 - W_2E_{C_3} - F_{B_3}W_3E_{C_4},$$

where Y_0 is a particular common solution of the pair $B_3YC_3 = A_3$ and $B_4YC_4 = A_4$, W_1—W_3 are arbitrary. Hence

$$\begin{aligned} X - Y &= X_0 - Y_0 + F_{B_2}V_1 + F_{B_4}W_1 + V_2E_{C_1} + W_2E_{C_3} + F_{B_1}V_3E_{C_2} + F_{B_3}W_3E_{C_4} \\ &= Z + [\, F_{B_2}, \ F_{B_4}\,] \begin{bmatrix} V_1 \\ W_1 \end{bmatrix} + [\, V_2, \ W_2\,] \begin{bmatrix} E_{C_1} \\ E_{C_3} \end{bmatrix} + F_{B_1}V_3E_{C_2} + F_{B_3}W_3E_{C_4}, \end{aligned}$$

where $Z = X_0 - Y_0$. Applying (27.15) to it, one can determine (28.25), we leave the routine work to the reader. Furthermore, we have the following useful consequence.

Theorem 28.9. *Suppose that the quadruple matrix equations* (28.24) *satisfy the condition* (28.26)*. Then they have a common solution if and only if the following fourteen rank equalities are all satisfied*

$$r[\,B_i,\ A_i\,] = r(B_i), \qquad r\begin{bmatrix} C_i \\ A_i \end{bmatrix} = r(C_i), \qquad i = 1,\,2,\,3,\,4, \tag{28.28}$$

$$r\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & -A_2 & B_2 \\ C_1 & C_2 & 0 \end{bmatrix} = r\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} + r[\,C_1,\ C_2\,], \tag{28.29}$$

$$r\begin{bmatrix} A_3 & 0 & B_3 \\ 0 & -A_4 & B_4 \\ C_3 & C_4 & 0 \end{bmatrix} = r\begin{bmatrix} B_3 \\ B_4 \end{bmatrix} + r[\,C_3,\ C_4\,], \tag{28.30}$$

$$r\begin{bmatrix} A_i & 0 & B_i \\ 0 & -A_j & B_j \\ C_i & C_j & 0 \end{bmatrix} = r\begin{bmatrix} B_i \\ B_j \end{bmatrix} + r[\,C_i,\ C_j\,], \qquad i = 1,\,2, \quad j = 3,\,4. \tag{28.31}$$

In fact, it is obvious that (28.24) has a common solution if and only if the two pairs $B_1XC_1 = A_1$, $B_2XC_2 = A_2$ and $B_3YC_3 = A_3$, $B_4YC_4 = A_4$ are consistent, respectively, and the equation

$$[\,F_{B_2},\ F_{B_4}\,]\begin{bmatrix} V_1 \\ W_1 \end{bmatrix} + [\,V_2,\ W_2\,]\begin{bmatrix} E_{C_1} \\ E_{C_3} \end{bmatrix} + F_{B_1}V_3E_{C_2} + F_{B_3}W_3E_{C_4} = Y_0 - X_0 \tag{28.32}$$

is consistent. According to corollary 20.3, the consistency conditions for the two pairs $B_1XC_1 = A_1$, $B_2XC_2 = A_2$ and $B_3YC_3 = A_3$, $B_4YC_4 = A_4$ are the ten rank equalities in (28.28)—(28.30). Next applying the four rank equalities in (27.17) and (27.18) to the equation (28.32) and simplifying, we can eventually find that (28.32) is consistent if and only if the four rank equalities in (28.31) hold. Based on Theorem 28.9, we are now able to establish a consistency condition for the matrix equation

$$B_1X_1C_1 + B_2X_2C_2 + B_3X_3C_3 = A, \tag{28.33}$$

which was presented in Theorem 27.9.

The Proof of Theorem 27.9. Write first (28.33) as

$$B_1X_1C_1 + B_2X_2C_2 = A - B_3X_3C_3. \tag{28.34}$$

Then by Corollary 27.3 we know that this equation is solvable if and only if there exists an X_3 satisfying the following four rank equalities

$$r[\,B_1,\ B_2,\ A - B_3X_3C_3\,] = r[\,B_1,\ B_2\,], \qquad r\begin{bmatrix} C_1 \\ C_2 \\ A - B_3X_3C_3 \end{bmatrix} = r\begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \tag{28.35}$$

$$r\begin{bmatrix} A - B_3X_3C_3 & B_1 \\ C_2 & 0 \end{bmatrix} = r(B_1) + r(C_2), \qquad r\begin{bmatrix} A - B_3X_3C_3 & B_2 \\ C_1 & 0 \end{bmatrix} = r(B_2) + r(C_1). \tag{28.36}$$

Applying (1.2)—(1.5) to the right hand sides of these four rank equalities, we see that they are equivalent to the following four matrix equations

$$E_PB_3X_3C_3 = E_PA, \qquad E_{B_1}B_3X_3C_3F_{C_2} = E_{B_1}AF_{C_2},$$
$$E_{B_2}B_3X_3C_3F_{C_1} = E_{B_2}AF_{C_1}, \qquad B_3X_3C_3F_Q = AF_Q,$$

where $P = [\,B_1,\ B_2\,]$ and $Q = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$, which can be simply written as

$$G_1XH_1 = L_1, \qquad G_2XH_2 = L_2, \qquad G_3XH_3 = L_3, \qquad G_34XH_4 = L_4, \tag{28.37}$$

where

$$G_1 = E_PB_3, \quad G_2 = E_{B_1}B_3, \quad G_3 = E_{B_2}B_3, \quad G_4 = B_3, \tag{28.38}$$
$$H_1 = C_3, \quad H_2 = C_3F_{C_2}, \quad H_3 = C_3F_{C_1}, \quad H_4 = C_3F_Q, \tag{28.39}$$
$$L_1 = E_PA, \quad L_2 = E_{B_1}CF_{B_2}, \quad L_3 = E_{B_2}AF_{C_1}, \quad L_4 = AF_Q. \tag{28.40}$$

It is not difficult to deduce that the given matrices in (28.37) satisfy the following four range inclusions

$$R(G_1^T) \subseteq R(G_2^T), \qquad R(H_2) \subseteq R(H_1), \qquad R(G_3^T) \subseteq R(G_4^T), \qquad R(H_4) \subseteq R(H_3). \tag{28.41}$$

Thus by Theorem 28.9 we know that the four equations in (28.37) have a common solution if and only if the following fourteen rank equalities all hold

$$r[\, G_i,\ L_i \,] = r(G_i), \qquad r\begin{bmatrix} H_i \\ L_i \end{bmatrix} = r(H_i), \qquad i = 1,\ 2,\ 3,\ 4, \tag{28.42}$$

$$r\begin{bmatrix} L_1 & 0 & G_1 \\ 0 & -L_i & G_i \\ H_1 & H_i & 0 \end{bmatrix} = r\begin{bmatrix} G_1 \\ G_i \end{bmatrix} + r[\, H_1,\ H_i \,], \qquad i = 2,\ 3,\ 4, \tag{28.43}$$

$$r\begin{bmatrix} L_i & 0 & G_i \\ 0 & -L_4 & G_4 \\ H_i & H_4 & 0 \end{bmatrix} = r\begin{bmatrix} G_i \\ G_4 \end{bmatrix} + r[\, H_i,\ H_4 \,], \qquad i = 2,\ 3, \tag{28.44}$$

$$r\begin{bmatrix} L_2 & 0 & G_2 \\ 0 & -L_3 & G_3 \\ H_2 & H_3 & 0 \end{bmatrix} = r\begin{bmatrix} G_2 \\ G_3 \end{bmatrix} + r[\, H_2,\ H_3 \,]. \tag{28.45}$$

Substituting the explicit expressions of $G_i,\ H_i$ and $L_i (i = 1,\ 2,\ 3,\ 4)$ into the eight rank equalities in (28.42) and simplifying by (1.2)—(1.4), we can find that they are equivalent to the first eight rank equalities in Theorem 27.9, respectively. Next substituting (28.38)—(28.40) into the five rank equalities in (28.43) and (28.44) and simplifying by (1.2)—(1.4), we can also find that they are equivalent to the first eight rank equalities in Theorem 27.9, respectively. We omit the routine processes here for simplicity. As for (28.45), we have by (1.2)—(1.4) that

$$\begin{aligned} r\begin{bmatrix} L_2 & 0 & G_2 \\ 0 & -L_3 & G_3 \\ H_2 & H_3 & 0 \end{bmatrix} &= r\begin{bmatrix} E_{B_1}AF_{C_2} & 0 & E_{B_1}B_3 \\ 0 & -E_{B_2}CF_{C_1} & E_{B_2}B_3 \\ C_3F_{C_2} & C_3F_{C_1} & 0 \end{bmatrix} \\ &= r\begin{bmatrix} A & 0 & B_1 & 0 & B_3 \\ 0 & -A & 0 & B_2 & B_3 \\ C_2 & 0 & 0 & 0 & 0 \\ 0 & C_1 & 0 & 0 & 0 \\ C_3 & C_3 & 0 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} - r\begin{bmatrix} C_2 & 0 \\ 0 & C_1 \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} r\begin{bmatrix} G_2 \\ G_3 \end{bmatrix} + r[\, H_2,\ H_3 \,] &= r\begin{bmatrix} E_{B_1}B_3 \\ E_{B_2}B_3 \end{bmatrix} + r[\, C_3F_{C_2},\ C_3F_{C_1} \,] \\ &= r\begin{bmatrix} B_1 & 0 & B_3 \\ 0 & B_2 & B_3 \end{bmatrix} + r\begin{bmatrix} C_2 & 0 \\ 0 & C_1 \\ C_3 & C_3 \end{bmatrix} - r\begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} - r\begin{bmatrix} C_2 & 0 \\ 0 & C_1 \end{bmatrix}. \end{aligned}$$

Thus (28.45) is equivalent to the last rank equality in Theorem 27.9. Summing up, we find that (28.37) has a common solution, or equivalently (28.33) is consistent, if and only if the nine rank equalities in Theorem 27.9 all hold. □

Chapter 29

Extreme ranks of $A - BX - XC$ subject to $BXC = D$

As a simple application of the rank formulas in Chapter 27, we determine in this chapter extreme ranks of a linear matrix equation $A - BX - XC$ subject to a consistent matrix $BXC = D$. This work is motivated by factoring a matrix D as $A = BB^{-} \pm B^{-}C$, and some related topics. Another motivation is from considering extreme ranks of $A - BX - XC$ subject to X. Quite different to the matrix expressions in the previous chapters, the same variant term X occurs two places in $A - BX - XC$. Although it is quite simple in form, we fail to establish a general method for expressing its extreme ranks except some special cases. An interesting exception is that when X is restricted by a consistent matrix equation $BXC = D$, extreme ranks of $A - BX - XC$ can completely be determined.

Theorem 29.1. *Let $A, D \in \mathcal{F}^{m \times n}$, $B \in \mathcal{F}^{m \times m}$ and $C \in \mathcal{F}^{n \times n}$ be given, and the matrix equation $BXC = D$ is consistent. Then*

$$\max_{BXC=D} r(\, A - BX - XC\,) = \min \left\{ m + r[\, BA - D,\ B^2\,] - r(B), \quad n + r\begin{bmatrix} AC - D \\ C^2 \end{bmatrix} - r(C), \right.$$
$$\left. r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}, \quad m + n + r(\, BAC - BD - DC\,) - r(B) - r(C) \right\}, \tag{29.1}$$

and

$$\min_{BXC=D} r(\, A - BX - XC\,) = r[\, BA - D,\ B^2\,] + r\begin{bmatrix} AC - D \\ C^2 \end{bmatrix} + \max\{\, s_1, \quad s_2\,\}, \tag{29.2}$$

where

$$s_1 = r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r\begin{bmatrix} C & 0 \\ BA & B^2 \end{bmatrix} - r\begin{bmatrix} B & AC \\ C^2 & 0 \end{bmatrix},$$
$$s_2 = r(\, BAC - BD - DC\,) - r[\, BAC - DC,\ B^2\,] - r\begin{bmatrix} BAC - BD \\ C^2 \end{bmatrix}.$$

Proof. Putting the general solution $X = B^{-}DC^{-} + F_BV_1 + V_2E_C$ of $BXC = D$ in $A - BX - XC$ we first get

$$A - BX - XC = A - DC^{-}B^{-}D - F_BV_1C - BV_2E_C = p(V_1,\, V_2), \tag{29.4}$$

Clearly, this is a linear matrix expression involving two independent variant matrices V_1 and V_2. In that case, we get by (27.6) and (27.7) that

$$\max_{V_1,\, V_2} r[\, p(V_1,\, V_2)\,] = \min \left\{ r[\, A_1,\, B,\, F_B,\,], \quad r\begin{bmatrix} A_1 \\ C \\ E_C \end{bmatrix}, \quad r\begin{bmatrix} A_1 & B \\ C & 0 \end{bmatrix}, \quad r\begin{bmatrix} A_1 & F_B \\ E_C & 0 \end{bmatrix} \right\}, \tag{29.5}$$

and

$$
\begin{aligned}
&\min_{V_1,\,V_2} r[\,p(V_1,\,V_2)\,] \\
&= r\begin{bmatrix} A_1 \\ C \\ E_C \end{bmatrix} + r[\,A,\,B,\,F_B\,] + \max\left\{ r\begin{bmatrix} A_1 & B \\ C & 0 \end{bmatrix} - r\begin{bmatrix} A_1 & B & F_B \\ C & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A_1 & B \\ C & 0 \\ E_C & 0 \end{bmatrix}, \right. \\
&\qquad\qquad \left. r\begin{bmatrix} A_1 & F_B \\ E_C & 0 \end{bmatrix} - r\begin{bmatrix} A_1 & E_B & B \\ E_C & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A_1 & E_B \\ E_C & 0 \\ C & 0 \end{bmatrix} \right\}, \qquad (29.6)
\end{aligned}
$$

where $A_1 = A - DC^- - B^-D$. Simplifying the ranks of the block matrix in them by Lemma 1.1, we have

$$
r[\,A_1,\,B,\,F_B\,] = r\begin{bmatrix} A - B^-D & B & I_m \\ 0 & 0 & B \end{bmatrix} - r(B) = r[\,BA - D,\;B^2\,] + m - r(B),
$$

$$
r\begin{bmatrix} A_1 \\ C \\ E_C \end{bmatrix} = r\begin{bmatrix} A - DC^- & 0 \\ C & 0 \\ I_n & C \end{bmatrix} - r(C) = r\begin{bmatrix} AC - D \\ C^2 \end{bmatrix} + n - r(C),
$$

$$
r\begin{bmatrix} A_1 & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A - DC^- - B^-D & B \\ C & 0 \end{bmatrix} = r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix},
$$

$$
\begin{aligned}
r\begin{bmatrix} A_1 & F_B \\ E_C & 0 \end{bmatrix} &= r\begin{bmatrix} A & I_m & 0 \\ I_n & 0 & C \\ 0 & B & 0 \end{bmatrix} - r(B) - r(C) \\
&= r\begin{bmatrix} 0 & I_m & 0 \\ I_n & 0 & 0 \\ 0 & 0 & BAC - BD - DC \end{bmatrix} - r(B) - r(C) \\
&= m + n + r(\,BAC - BD - DC\,) - r(B) - r(C),
\end{aligned}
$$

$$
\begin{aligned}
r\begin{bmatrix} A_1 & B & F_B \\ E_C & 0 & 0 \end{bmatrix} &= r\begin{bmatrix} A & B & I_m \\ C & 0 & 0 \\ 0 & 0 & B \end{bmatrix} - r(B) \\
&= r\begin{bmatrix} 0 & 0 & I_m \\ C & 0 & 0 \\ BA & B^2 & 0 \end{bmatrix} - r(B) = m + r\begin{bmatrix} C & 0 \\ BA & B^2 \end{bmatrix} - r(B),
\end{aligned}
$$

$$
\begin{aligned}
r\begin{bmatrix} A_1 & B \\ C & 0 \\ E_C & 0 \end{bmatrix} &= r\begin{bmatrix} A & B & 0 \\ C & 0 & 0 \\ I_n & 0 & C \end{bmatrix} - r(C) \\
&= r\begin{bmatrix} 0 & B & AC \\ 0 & 0 & C^2 \\ I_n & 0 & 0 \end{bmatrix} - r(C) = n + r\begin{bmatrix} B & AC \\ 0 & C^2 \end{bmatrix} - r(C),
\end{aligned}
$$

$$
\begin{aligned}
r\begin{bmatrix} A_1 & F_B & B \\ E_C & 0 & 0 \end{bmatrix} &= r\begin{bmatrix} A - B^-D & I_m & B & 0 \\ I_n & 0 & 0 & C \\ 0 & B & 0 & 0 \end{bmatrix} - r(B) - r(C) \\
&= r\begin{bmatrix} 0 & I_m & 0 & 0 \\ I_n & 0 & 0 & 0 \\ 0 & 0 & B^2 & BAC - DC \end{bmatrix} - r(B) - r(C) \\
&= r[\,B^2,\;BAC - DC\,] - r(B) - r(C) + m + n,
\end{aligned}
$$

$$
r\begin{bmatrix} A_1 & F_B \\ E_C & 0 \\ C & 0 \end{bmatrix} = r\begin{bmatrix} ADC^- & I_m & 0 \\ I_n & 0 & C \\ C & 0 & 0 \\ 0 & B & 0 \end{bmatrix} - r(B) - r(C)
$$

$$= r\begin{bmatrix} 0 & I_m & 0 \\ I_n & 0 & 0 \\ 0 & 0 & C^2 \\ 0 & 0 & BAC - BD \end{bmatrix} - r(B) - r(C)$$
$$= r\begin{bmatrix} C^2 \\ BAC - BD \end{bmatrix} - r(B) - r(C) + m + n.$$

Putting them in (29.3) and (29.4) yields (29.1) and (29.2). □

Corollary 29.2. *Let $A,\, D \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times m}$ and $C \in \mathcal{F}^{n\times n}$ be given, and the two matrix equations $BX + YC = A$ and $BXC = D$ are consistent, respectively. Then*

$$\max_{BXC=D} r(\, A - BX - XC\,) = \min\Big\{ m + r[\, BA - D,\; B^2\,] - r(B), \quad n + r\begin{bmatrix} AC - D \\ C^2 \end{bmatrix} - r(C),$$
$$r(B) + r(C), \quad m + n + r(\, BAC - BD - DC\,) - r(B) - r(C) \Big\}, \tag{29.7}$$

and

$$\min_{BXC=D} r(\, A - BX - XC\,) = r[\, BA - D,\; B^2\,] + r\begin{bmatrix} AC - D \\ C^2 \end{bmatrix}$$
$$+ \max\Big\{ -r(B^2) - r(C^2), \quad r(\, BAC - BD - DC\,) - r[\, BAC - DC,\; B^2\,] - r\begin{bmatrix} BAC - BD \\ C^2 \end{bmatrix} \Big\}. \tag{29.8}$$

Proof. The consistency of $BX + YC = A$ implies that $r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(B) + r(C)$. Thus (29.1) and (29.2) reduce to (29.7) and (29.8) □

Corollary 29.3. *Let $A,\, D \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times m}$ and $C \in \mathcal{F}^{n\times n}$ be given, and $BXC = D$ is consistent. If $r(B^2) = r(B)$ and $r(C^2) = r(C)$, then*

$$\max_{BXC=D} r(\, A - BX - XC\,) = \min\Big\{ m + r[\, D,\; B\,] - r(B), \quad n + r\begin{bmatrix} D \\ C \end{bmatrix} - r(C),$$
$$r(B) + r(C), \quad m + n + r(\, BAC - BD - DC\,) - r(B) - r(C) \Big\}, \tag{29.9}$$

and

$$\min_{BXC=D} r(\, A - BX - XC\,) = r[\, D,\; B\,] + r\begin{bmatrix} D \\ C \end{bmatrix} + r(\, BAC - BD - DC\,) - r(B) - r(C). \tag{29.10}$$

Proof. Under $r(B^2) = r(B)$ and $r(C^2) = r(C)$, there are

$$r[\, BA - D,\; B^2\,] = r[\, D,\; B\,], \quad r\begin{bmatrix} AC - D \\ C^2 \end{bmatrix} = r\begin{bmatrix} D \\ C \end{bmatrix}, \quad r[\, BAC - DC,\; B^2\,] = r(B),$$
$$r\begin{bmatrix} BAC - BD \\ C^2 \end{bmatrix} = r(C), \quad \begin{bmatrix} BA & B^2 \\ C & 0 \end{bmatrix} = \begin{bmatrix} B & AC \\ C^2 & 0 \end{bmatrix} = r(B) + r(C).$$

Thus we have (29.9) and (29.10). □

Corollary 29.4. *Let $A,\, D \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times m}$ and $C \in \mathcal{F}^{n\times n}$ be given, and the two matrix equations $BX + YC = A$ and $BXC = D$ are consistent, respectively. Then the pair of matrix equations*

$$BX + XC = A \qquad BXC = D \tag{29.11}$$

have a common solution if and only if the following three conditions hold

$$R(\, BA - D\,) \subseteq R(B^2), \qquad R[(\, AC - D\,)^T] \subseteq R[(C^2)^T], \qquad BD + DC = BAC. \tag{29.12}$$

Proof. Letting the right hand side of (29.8) be zero and simplifying yield (29.12). □

If $BX + XC = A$ and $BXC = D$ have a common solution, their general common solution can be simply found by the following two steps: First solve the matrix equation

$$F_BV_1C + BV_2E_C = A - DC^- - B^-D \tag{29.13}$$

for V_1 and V_2. Then put V_1 and V_2 in $X = B^-DC^- + F_BV_1 + V_2E_C$ to yield the general common solution to the pair of equations. Based on the results in Lemma 26.2, we find that their general common solution can be written as

$$X = X_0 + [\, F_B,\ 0\,]F_GUE_H \begin{bmatrix} I_n \\ 0 \end{bmatrix} + [\, 0,\ I_m\,]F_GUE_H \begin{bmatrix} 0 \\ E_C \end{bmatrix} + F_BSE_C, \tag{29.14}$$

where X_0 is a particular common solution to (29.11), $G = [\, F_B,\ -B\,]$, $H = \begin{bmatrix} C \\ E_C \end{bmatrix}$, U and S are arbitrary. From Corollary 29.4 and (29.14) we also see that (29.11) has a unique common solution if and only if both B and C are nonsingular and $BD + DC = BDC$. In that case, the unique common solution is $X = B^{-1}DC^{-1}$.

Theorem 29.1 can apply to determine extreme ranks of the matrix expressions $A - BB^- - B^-B$ and $A - BB^- + B^-B$ with respect to B^-. In fact

$$\max_{B^-} r(\, A - BB^- - B^-B\,) = \max_{BXB=B} r(\, A - BX - XB\,),$$
$$\min_{B^-} r(\, A - BB^- - B^-B\,) = \min_{BXB=B} r(\, A - BX - XB\,),$$
$$\max_{B^-} r(\, A - BB^- + B^-B\,) = \max_{BX(-B)=-B} r[\, A - BX - X(-B)\,],$$
$$\min_{B^-} r(\, A - BB^- + B^-B\,) = \min_{BX(-B)=-B} r[\, A - BX - X(-B)\,].$$

Applying Theorem 29.1, one can easily get the maximal and the minimal ranks of the matrix expressions, as well as necessary and sufficient conditions for the factorization $A = BB^- + B^-B$ or $A = BB^- - B^-B$ to hold.

As we mentioned in beginning the chapter, it is a quite difficult problem to find in general extreme ranks of the linear matrix expressions $A - BX - XC$ as well as $A - B_1XC_1 - B_2XC_2$. However, if the given matrices in them satisfy conditions, we can find their extreme ranks. Here we present two special results related to the minimal ranks of $A - BX + XC$ and $A - X + BXC$ when both B and C are idempotent.

Theorem 29.5. *Suppose that B and C are $m \times m$ and $n \times n$ idempotent matrices, respectively. Then*

$$\min_X r(\, A - BX + XC\,) = \max\{\, r(BAC), \quad r(\, I_m - B\,)A(\, I_n - C\,)\,\}, \tag{29.15}$$

$$\min_X r(\, A - X + BXC\,) = r(BAC). \tag{29.16}$$

In particular,

(a) *The matrix equation $BX - XC = A$ is consistent if and only if $ABC = 0$ and $r(\, I_m - B\,)A(\, I_n - C\,) = 0$.*

(b) *The matrix equation $X - BXC = A$ is consistent if and only if $ABC = 0$.*

Proof. Observe that

$$B(\, A - X + BXC\,)C = BAC - BXC + B^2XC^2 = BAC.$$

We first see that $r(\, A - X + BXC\,) \geq r(BAC)$ holds for all X. On the other hand, let $X = A$, then $A - A + BAC = BAC$. The combination of the above two facts yields (29.16).

To prove (29.15), we use the simple result

$$\min_{X,\, Y} r \begin{bmatrix} M & X \\ Y & N \end{bmatrix} = \max\{\, r(M), \quad r(N)\,\}. \tag{29.17}$$

Since both B and C are idempotent, we can factor them as

$$B = P^{-1} \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} P, \quad C = Q \begin{bmatrix} I_l & 0 \\ 0 & 0 \end{bmatrix} Q^{-1},$$

where $k = r(B)$ and $l = r(C)$. In that case,

$$A - BX + XC = P^{-1}\left(PAQ - \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} PXQ + PXQ \begin{bmatrix} I_l & 0 \\ 0 & 0 \end{bmatrix} \right) Q^{-1}. \tag{29.18}$$

Let $Y = PXQ = \begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix}$ and $PAQ = \begin{bmatrix} S_1 & S_2 \\ S_3 & S_4 \end{bmatrix}$. Then from (29.18) we get

$$r(\, A - BX + XC\,) = r\left(\begin{bmatrix} S_1 & S_2 \\ S_3 & S_4 \end{bmatrix} - \begin{bmatrix} Y_1 & Y_2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} Y_1 & 0 \\ Y_3 & 0 \end{bmatrix} \right) = r \begin{bmatrix} S_1 & S_2 - Y_2 \\ S_3 + Y_3 & S_4 \end{bmatrix}.$$

Applying (29.17) to it we find

$$\min_X r(\, A - BX + XC\,) = \min_{Y_2,\, Y_3} r \begin{bmatrix} S_1 & S_2 - Y_2 \\ S_3 + Y_3 & S_4 \end{bmatrix} = \min\{\, r(S_1), \;\; r(S_4)\,\},$$

where

$$r(S_1) = r\left(\begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} PAQ \begin{bmatrix} I_l & 0 \\ 0 & 0 \end{bmatrix} \right) = r\left(P^{-1} \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} PAQ \begin{bmatrix} I_l & 0 \\ 0 & 0 \end{bmatrix} Q^{-1} \right) = r(BAC).$$

and

$$\begin{aligned} r(S_4) &= r\left(\begin{bmatrix} 0 & 0 \\ 0 & I_{m-k} \end{bmatrix} PAQ \begin{bmatrix} 0 & 0 \\ 0 & I_{n-l} \end{bmatrix} \right) \\ &= r\left(P^{-1} \begin{bmatrix} 0 & 0 \\ 0 & I_{m-k} \end{bmatrix} PAQ \begin{bmatrix} 0 & 0 \\ 0 & I_{n-l} \end{bmatrix} Q^{-1} \right) = r(\, I_m - B\,)A(\, I_n - C\,). \end{aligned}$$

Thus we have (29.15). $\square$

Finally we present another interesting result related to the minimal rank of a complex matrix with respect to its imaginary part. It could also be regarded a special case of a matrix expression $A - B_1XC_1 - B_2XC_2$. We leave it as an exercise to the reader.

Theorem 29.6. *Suppose that A and X are two real matrices of the same size. Then*

$$\min_X r(\, A + iX\,) = \frac{1}{2} \min_X r \begin{bmatrix} A & -X \\ X & A \end{bmatrix} = \begin{cases} r(A)/2 & if \ r(A) \ is \ even \\ [\, r(A) + 1\,]/2 & if \ r(A) \ is \ odd \end{cases}.$$

The problem can also reasonably be considered for a quaternion matrix. Here we list a conjecture.

Conjecture 29.7. *Let $A + iX + jY + kZ$ be a quaternion matrix, where $i^2 = j^2 = k^2 = -1$ and $ijk = -1$, the matrices A, X_1, X_2 and X_3 are real. Then*

$$\min_{X,\,Y,\,Z} r(\, A + iX + jY + kZ\,) = \begin{cases} r(A)/4 & if \ r(A) \ \equiv 0 \ (\mathrm{mod} = 4) \\ [\, r(A) + 3\,]/4 & if \ r(A) \ \equiv 1 \ (\mathrm{mod} = 4) \\ [\, r(A) + 2\,]/4 & if \ r(A) \ \equiv 2 \ (\mathrm{mod} = 4) \\ [\, r(A) + 1\,]/4 & if \ r(A) \ \equiv 3 \ (\mathrm{mod} = 4). \end{cases}$$

Chapter 30

Extreme ranks of some quadratic matrix expressions

Without much effort, the work in previous chapters can be easily extended quadratic matrix expressions involving two independent variant matrices. In this chapter we first present the maximal and the minimal ranks of a matrix expression

$$q(X_1,\, X_2) = A - (\, A_1 - B_1X_1C_1\,)D(\, A_2 - B_2X_2C_2\,) \tag{30.1}$$

subject to X_1 and X_2, and then present their various consequences. The fundamental tool used for coping with (30.1) is the following rank formula

$$r(\, A - PNQ\,) = r\begin{bmatrix} A & PN \\ NQ & N \end{bmatrix} - r(N). \tag{30.2}$$

Applying (30.2) to (30.1), we can get

$$\begin{aligned} & r[q(X_1,\, X_2)] \\ & = r\left(\begin{bmatrix} A & (\, A_1 - B_1X_1C_1\,)D \\ D(\, A_2 - B_2X_2C_2\,) & D \end{bmatrix}\right) - r(D) \\ & = r\left(\begin{bmatrix} A & A_1D \\ DA_2 & D \end{bmatrix} - \begin{bmatrix} B_1 \\ 0 \end{bmatrix} X_1[\,0,\; C_1D\,] - \begin{bmatrix} 0 \\ DB_2 \end{bmatrix} X_1[\,C_2,\; 0\,]\right) - r(D). \end{aligned} \tag{30.3}$$

Evidently the matrix expression in the right hand side of (30.3) is linear with two independent variant matrices. Applying the rank formulas (27.6) and (27.7) to (30.3) and simplifying, we get the following.

Theorem 30.1. *Let $q(X_1,\, X_2)$ be given by* (30.1)*. Then*

$$\begin{aligned} \max_{X_1,\, X_2} r[\,q(X_1,\, X_2)\,] = \min\Bigg\{ & r[\,A - A_1DA_2,\; A_1DB_2,\; B_1\,],\;\; r\begin{bmatrix} A - A_1DA_2 \\ C_1DA_2 \\ C_2 \end{bmatrix}, \\ & r\begin{bmatrix} A_1DA_2 - A & B_1 \\ C_2 & 0 \end{bmatrix},\;\; r\begin{bmatrix} A_1DA_2 - A & A_1DB_2 \\ C_1DA_2 & C_1DB_2 \end{bmatrix} \Bigg\}, \end{aligned} \tag{30.4}$$

and

$$\begin{aligned} \min_{X_1,\, X_2} r[\,q(X_1,\, X_2)\,] = {} & r\begin{bmatrix} A - A_1DA_2 \\ C_1DA_2 \\ C_2 \end{bmatrix} + r[\,A - A_1DA_2,\; A_1DB_2,\; B_1\,] \\ & + \max\Bigg\{ r\begin{bmatrix} A_1DA_2 - A & B_1 \\ C_2 & 0 \end{bmatrix} - r\begin{bmatrix} A_1DA_2 - A & B_1 & A_1DB_2 \\ C_2 & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A_1DA_2 - A & B_1 \\ C_2 & 0 \\ C_1DA_2 & 0 \end{bmatrix}, \end{aligned}$$

$$r\begin{bmatrix} A_1DA_2 - A & A_1DB_2 \\ C_1DA_2 & C_1DB_2 \end{bmatrix} - r\begin{bmatrix} A_1DA_2 - A & A_1DB_2 & B_1 \\ C_1DA_2 & C_1DB_2 & 0 \end{bmatrix} - r\begin{bmatrix} A_1DA_2 - A & A_1DB_2 \\ C_1DA_2 & C_1DB_2 \\ C_2 & 0 \end{bmatrix}\Bigg\}. \quad (30.5)$$

The formulas (30.4) and (30.5) can further simplify when the given matrices in them satisfy some conditions, for example, the two equations $B_1X_1C_1 = A_1$ and $B_2X_2C_2 = A_2$ are solvable, respectively; or some of them are identity matrices or zero matrices.

Two nice results are given below.

Corollary 30.2. *Let $q(X, Y) = XAY + XB + CY + D$, where where A, B, C and D are $m \times n$, $l \times n$, $m \times k$, and $l \times k$ matrices, respectively. Then*

(a) *The maximal and the minimal ranks of $q(X_1, X_2)$ are*

$$\max_{X,\,Y} r[\,q(X,\,Y)\,] = \min\left\{ m, \quad n, \quad r\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right\}, \quad (30.6)$$

$$\min_{X,\,Y} r[\,q(X,\,Y)\,] = \max\left\{ 0, \quad r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A \\ C \end{bmatrix} - r[\,A,\ B\,]\right\}. \quad (30.7)$$

(b) *Let $m = n$. Then there are X and Y such that $q(X, Y)$ is nonsingular if and only if*

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} \geq m. \quad (30.8)$$

(c) *There are X and Y such that $XAY + XB + YC + D = 0$ if and only if*

$$r\begin{bmatrix} A & B \\ C & D \end{bmatrix} \leq r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,]. \quad (30.9)$$

The matrix expression $q(X, Y) = XAY + XB + CY + D$ occurs in an elementary operation for a 2×2 block matrix

$$\begin{bmatrix} I_m & 0 \\ X & I_l \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} I_n & Y \\ 0 & I_k \end{bmatrix} = \begin{bmatrix} A & AY + B \\ C + XA & XAY + XB + CY + D \end{bmatrix}. \quad (30.10)$$

Clearly the lower right block in (30.10) is the matrix expression $q(X, Y)$. If we let $X = -CA^-$ and $Y = -A^-B$, where A^- is an inner inverse of A, then $q(X, Y) = D - CA^-B$, the well-known Schur complement A in $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$. Corollary 30.2 acctually gives possible ranks of the lower right block of (30.10) after the block elementary operation, including necessary and sufficient conditions for the block to be nonsingular or null.

Corollary 30.3. *Let $q(X_1, X_2) = A - BX_1DX_2C$, where A, B, C and D are $m \times n$, $m \times k$, $l \times n$, and $p \times q$ matrices, respectively. Then the maximal and the minimal ranks of $q(X_1, X_2)$ are*

$$\max_{X_1,\,X_2} r[\,q(X_1,\,X_2)\,] = \min\left\{ r\begin{bmatrix} A \\ C \end{bmatrix}, \quad r[\,A,\ B\,], \quad r(A) + r(D)\right\}, \quad (30.10)$$

$$\min_{X_1,\,X_2} r[\,q(X_1,\,X_2)\,] = \max\left\{ r(A) - r(D), \quad r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,] - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\right\}. \quad (30.11)$$

Theoretically one can express rank of any nonlinear matrix expression through rank of a linear matrix expression. A basic transformation formula for this is

$$r(\,A - B_1X_1B_2X_2 \cdots B_kX_kB_{k+1}\,)$$
$$= r\begin{bmatrix} B_1X_1B_2 & 0 & \cdots & 0 & (-1)^kA \\ B_2 & B_2X_2B_3 & \cdots & 0 & 0 \\ & \ddots & \ddots & \vdots & \vdots \\ & & & B_{k-1}X_{k-1}B_k & 0 \\ & & & B_k & B_kX_kB_{k+1} \end{bmatrix} - r(B_2) - \cdots - r(B_k). \quad (30.12)$$

The block matrix on the right side of (30.12) is obviously a linear matrix expression.

As an important application we next consider extreme ranks of the Schur complement $D - CA_r^- B$ with respect to an reflexive inner inverse A_r^- of A. A reflexive inner inverse of A is a solution of the pair of matrix equations $AXA = A$ and $XAX = X$. The general expression of reflexive inner inverse of A can be written as $A_r^- = A^- A A^- = (\, A^\sim - F_A V_1 \,) A (\, A^\sim - V_2 E_A \,)$, where $A^\sim$ is a particular inner inverse of A, V_1 and V_2 are arbitrary. Therefore, we have

$$D - CA_r^- B = D - (\, CA^\sim - CF_A V_1 \,) A (\, A^\sim B - V_2 E_A B \,),$$

Applying (30.4) and (30.5) to it we get the following.

Theorem 30.4. *The maximal and the minimal ranks of the Schur complement $D - CA_r^- B$ with respect to A_r^- are given by*

$$\max_{A_r^-} r(\, D - CA_r^- B \,) = \min\left\{ r(A) + r(D), \quad r[\, A, \ B \,], \quad r\begin{bmatrix} A \\ C \end{bmatrix}, \quad r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r(A) \right\}, \tag{30.13}$$

and

$$\min_{A_r^-} r(\, D - CA_r^- B \,) = r\begin{bmatrix} B \\ D \end{bmatrix} + r[\, C, \ D \,] + r(A)$$

$$+ \max\left\{ r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix}, \ r(D) - r\begin{bmatrix} A & 0 \\ C & D \end{bmatrix} - r\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \right\}. \tag{30.14}$$

Remark. Just as what we did in Chapters 21, 22 and 23, one can further find many consequences from (30.13) and (30.14), such as, the rank invariance of $D - CA_r^- B$ with respect to the choice of A_r^-; various special cases of (30.13) and (30.14) and their interesting consequences when A, B, C and D satisfy some conditions; reverse order laws for reflexive inner inverses of products of matrices, rank equalities for sums of reflexive inner inverses of matrices, and so on. The reader can easily list them and apply them to find some more interesting results.

Through (30.4) and (30.5) we cal also derive various rank equalities for matrix expressions involving products of inner inverses of matrices. We next list several of them without detailed proofs.

Theorem 30.5. *Let $A \in \mathcal{F}^{m \times n}$, $B \in \mathcal{F}^{m \times k}$ and $C \in \mathcal{F}^{l \times n}$ be given. Then*

$$\begin{aligned} &\max_{B^-,\, C^-} r[\, A - (\, I_m - BB^- \,) A (\, I_n - C^- C \,) \,] \\ &= \min\left\{ r(B) + r(C), \quad r[\, A, \ B \,], \quad r\begin{bmatrix} A \\ C \end{bmatrix}, \quad r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} + r(A) - r(B) - r(C) \right\}, \end{aligned} \tag{30.15}$$

and

$$\min_{B^-,\, C^-} r[\, A - (\, I_m - BB^- \,) A (\, I_n - C^- C \,) \,] = r(A) + r(B) + r(C) - \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}. \tag{30.16}$$

Proof. Notice the two general expressions $BB^- = BB^\sim + F_B V_1 B$ and $C^- C = CC^\sim + CV_2 E_C$, where $B^\sim$ and $C^\sim$ are two particular inner inverses of B and C, respectively, V_1 and V_2 are arbitrary. Put them in $A - (\, I_m - BB^- \,) A (\, I_n - C^- C \,)$ to yield a quadratic matrix expression. In this case applying (30.4) and (30.5) to it and then simplifying we may trivially give (30.15) and (30.16). $\square$

Theorem 30.6. *Let $A \in \mathcal{F}^{m \times n}$, $B \in \mathcal{F}^{p \times m}$ and $C \in \mathcal{F}^{n \times q}$ be given. Then*

$$\max_{B^-,\, C^-} r[\, (\, I_m - B^- B \,) A (\, I_n - CC^- \,) \,] = \min\{\, r(A), \quad m - r(B), \quad n - r(C) \,\}, \tag{30.17}$$

$$\min_{B^-,\, C^-} r[\, (\, I_m - B^- B \,) A (\, I_n - CC^- \,) \,] = \max\{\, 0, \quad r(A) - r(BA) - r(AC) \,\}. \tag{30.18}$$

In particular, there are B^- and C^- such that $(\, I_m - B^- B \,) A (\, I_n - CC^- \,) = 0$, if and only if $r(A) \leq r(BA) + r(AC)$.

Some special cases of (30.17) and (30.18) are listed below:

$$\max_{B^-,\,C^-} r[\,(\,I_m - B^-B\,)(\,I_m - CC^-\,)\,] = \min\{\,m - r(B), \quad m - r(C)\,\},$$
$$\min_{B^-,\,C^-} r[\,(\,I_m - B^-B\,)(\,I_m - CC^-\,)\,] = \max\{\,0, \quad m - r(B) - r(C)\,\},$$
$$\max_{A^-} r[\,(\,I_m - A^-A\,)(\,I_m - AA^-\,)\,] = m - r(A),$$
$$\min_{A^-} r[\,(\,I_m - A^-A\,)(\,I_m - AA^-\,)\,] = \max\{\,0, \quad m - 2r(A)\,\}.$$

Theorem 30.7. *Let $A \in \mathcal{F}^{m\times m}$ be given. Then*

$$\max_{A^-} r[\,A - (\,I_m - AA^-\,)(\,I_m - A^-A\,)\,] = \min_{A^-} r[\,A - (\,I_m - AA^-\,)(\,I_m - A^-A\,)\,] = m + r(A^2) - r(A). \tag{30.19}$$

Theorem 30.8. *Let $A \in \mathcal{F}^{m\times m}$ be given. Then*

$$\max_{A^-} r[\,A - (\,I_m - A^-A\,)(\,I_m - AA^-\,)\,] = m + r(A^2) - r(A), \tag{30.20}$$
$$\min_{A^-} r[\,A - (\,I_m - A^-A\,)(\,I_m - AA^-\,)\,] = \max\{\,2r(A^2) - r(A^3), \quad m - 2r(A) + 2r(A^2) - r(A^3)\,\}. \tag{30.21}$$

Theorem 30.9. *Let $A,\ D \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ and $C \in \mathcal{F}^{l\times n}$ be given. Then*

$$\max_{B^-,\,C^-} r(\,A - BB^-DC^-C\,) = \min\left\{ r[\,A,\ B\,], \quad r\begin{bmatrix} A \\ C \end{bmatrix}, \quad r\begin{bmatrix} A & 0 & B \\ 0 & -D & B \\ C & C & 0 \end{bmatrix} - r(B) - r(C) \right\}, \tag{30.22}$$

and

$$\min_{B^-,\,C^-} r(\,A - BB^-DC^-C\,) = \min\{\,s_1, \quad s_2\,\}, \tag{30.23}$$

where

$$s_1 = r[\,A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix},$$

$$s_2 = r(B) + r(C) + r[\,A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} + r\begin{bmatrix} A & 0 & B \\ 0 & -D & B \\ C & C & 0 \end{bmatrix} - r\begin{bmatrix} A & 0 & B & 0 \\ 0 & D & 0 & B \\ C & C & 0 & 0 \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & D & B \\ C & 0 & 0 \\ 0 & C & 0 \end{bmatrix}.$$

Some useful consequences of (30.22) and (30.23) are listed below:

Theorem 30.10. *Let $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$ and $C \in \mathcal{F}^{l\times n}$ be given. Then*

$$\max_{B^-,\,C^-} r(\,A - BB^-AC^-C\,) = \min\left\{ r\begin{bmatrix} A \\ C \end{bmatrix}, \quad r[\,A,\ B\,], \quad r\begin{bmatrix} A \\ C \end{bmatrix} + r[\,A,\ B\,] - r(B) - r(C) \right\}, \tag{30.24}$$

and

$$\min_{B^-,\,C^-} r(\,A - BB^-AC^-C\,) = r[\,A,\ B\,] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}. \tag{30.25}$$

Contrasting (30.25) with (18.6), we see that

$$\min_{B^-,\,C^-} r(\,A - BB^-AC^-C\,) = \min_X r(\,A - BXC\,),$$

which can be stated that there is a matrix X with form $X = B^-AC^-$ such that $A - BXC$ reaches to its minimal rank. We can call this X a minimal rank solution of $BXC = A$. In that case, $X = B^-AC^- + F_AV_1 + V_2E_B$ is also minimizing $r(\,A - BXC\,)$, where V_1 and V_2 are arbitrary.

Theorem 30.11. *Let $B,\,C \in \mathcal{F}^{m\times m}$ be given. Then*

$$\max_{B^-,\,C^-} r(\,BC - BB^-C^-C\,) = \min\{\,r(B), \quad r(C), \quad m + r(\,C - CBC\,) - r(C)\,\}, \tag{30.26}$$

$$\min_{B^-,\,C^-} r(\, BC - BB^-C^-C\,) = \max\{\,0, \quad r(\,C - CBC\,) + r(B) - m\,\},$$

and (30.27)

$$\max_{B^-} r(\, B^2 - BB^-B^-B\,) = \min\{\, r(B), \quad m + r(\,B - B^3\,) - r(B)\,\}, \tag{30.28}$$

$$\min_{B^-} r(\, B^2 - BB^-B^-B\,) = \max\{\,0, \quad r(\,B - B^3\,) + r(B) - m\,\}. \tag{30.29}$$

In particular, there are B^- and C^- such that BC can factor as $BC = (BB^-)(C^-C)$ if and only if $r(\,C - CBC\,) \leq m - r(B)$. There is B^- such that B^2 can factor as $B^2 = (BB^-)(B^-B)$ if and only if $r(\,B - B^3\,) \leq m - r(B)$.

Theorem 30.12. *Let $A \in \mathcal{F}^{m\times m}$ be given. Then*

$$\max_{A^-} r(\, A - AA^-A^-A\,) = \max\{\, r(A), \quad r(\,I_m - A\,)\,\}, \tag{30.30}$$

$$\min_{A^-} r(\, A - AA^-A^-A\,) = \max\{\,0, \quad r(\,I_m - A\,) + 2r(A) - 2m\,\}. \tag{30.31}$$

In particular, there is an A^- such that A can factor as $A = (AA^-)(A^-A)$ if and only if $r(\,I_m - A\,) \leq 2m - 2r(A)$. Moreover, when $r(A) \leq m/2$, there must exist an A^- such that A can factor as $A = (AA^-)(A^-A)$.

Theorem 30.13. *Let $B,\, C \in \mathcal{F}^{m\times m}$ be given. Then*

$$\max_{B^-,\,C^-} r(\, BB^-C^-C\,) = \max\{\, r(B), \quad r(C)\,\}, \tag{30.32}$$

$$\min_{B^-,\,C^-} r(\, B^-C^-\,) = \min_{B^-,\,C^-} r(\, BB^-C^-C\,) = \max\{\,0, \quad r(B) + r(C) - m\,\}. \tag{30.33}$$

In particular, there are B^- and C^- such that $B^-C^- = 0$ if and only if $r(B) + r(C) \leq m$.

Moreover one can also find extreme ranks of matrix expressions $A - (\,I_m - BB^-\,)D(\,I_n - C^-C\,)$, $A - (\,I_m - B^-B\,)D(\,I_n - CC^-\,)$, $A - B^-BDCC^-$, $A - B^kB^-DC^-C^k$, and so on. Based on them more consequences can be derived. We leave them to the reader. In addition, we present another interesting result for the reader to prove

$$\min_{A^-,\,B^-,\,C^-} r(\, AA^-BB^-CC^-\,) = \dim[\, R(A) \cap R(B) \cap R(C)\,]. \tag{30.34}$$

Finally we present a conjecture on the minimal rank of multiple product of inner inverses.

Conjecture 30.14. *Let $A_i \in \mathcal{F}^{m_{i+1}\times m_i}$, $i = 1,\, 2,\, \cdots,\, k$. Then*

$$\min_{A_1^-,\,\cdots\, A_k^-} r(\, A_1^-A_2^- \cdots A_k^-\,) = \max\{\,0, \quad r(A_1) + r(A_2) + \cdots + r(A_k) - m_2 - m_3 - \cdots - m_k\,\}. \tag{30.35}$$

Chapter 31

Completing triangular block matrices with extreme ranks

Suppose that A_n and X_n are triangular block matrices with the forms

$$A_n = \begin{bmatrix} A_{11} & & & \\ A_{21} & A_{22} & & \\ \vdots & \vdots & \ddots & \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}, \qquad X_n = \begin{bmatrix} 0 & X_{12} & \cdots & X_{1n} \\ & \ddots & \ddots & \vdots \\ & & 0 & X_{n-1,n} \\ & & & 0 \end{bmatrix}, \tag{31.1}$$

where A_{ij} $(n \geq i \geq j \geq 1)$ is a given $s_i \times t_j$ matrix, X_{ij} $(1 \leq i < j \leq n)$ is a variant $s_i \times t_j$ matrix. Further let $S(X_n)$ be the collection of all matrices X_n in (31.1). In this article we consider how to choose $X_n \in S(X_n)$ such that

$$r(\, A_n + X_n \,) = \begin{bmatrix} A_{11} & X_{12} & \cdots & X_{1n} \\ A_{21} & A_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & X_{n-1,n} \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} = \max \tag{31.2}$$

and

$$r(\, A_n + X_n \,) = \begin{bmatrix} A_{11} & X_{12} & \cdots & X_{1n} \\ A_{21} & A_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & X_{n-1,n} \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} = \min \tag{31.3}$$

hold, respectively.

These two problems are well known in matrix theory as maximal and minimal rank completion problems, which have been previously examined by lots of authors from different aspects (see, e.g., [37, 47, 75, 76, 152, 153]). In this chapter, we wish to give a new investigation to the two problems by making use of the theory of generalized inverses of matrices.

Lemma 31.1. *Suppose that*

$$M(X_{12}) = \begin{bmatrix} A_{11} & X_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{31.4}$$

is a 2×2 block matrix, where A_{11}, A_{21} and A_{22} are three given $s_1 \times t_1, s_2 \times t_1$ and $s_2 \times t_2$ matrices, respectively, and X_{12} is a variant $s_1 \times t_2$ matrix. Then

(a) *The maximal rank of $M(X_{12})$ with respect to X_{12} is*

$$\max_{X_{12}} r[M(X_{12})] = \min \left\{ r \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} + t_2, \quad r[\, A_{21}, \;\, A_{22} \,] + s_1 \right\}, \tag{31.5}$$

and the matrix X_{12} *satisfying* (31.5) *can be expressed as*

$$X_{12} = \widehat{X}_{12} + A_{11}A_{21}^{-}A_{22} + A_{11}F_{A_{21}}V + WE_{A_{21}}A_{22}, \tag{31.6}$$

where V *and* W *are two arbitrary matrices,* $\widehat{X}_{12}$ *is chosen such that*

$$\begin{aligned} r(E_G\widehat{X}_{12}F_H) &= \min\{\, r(E_G), \quad r(F_H)\,\} \\ &= \min\left\{ s_1 + r(A_{21}) - r\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}, \quad t_2 + r(A_{21}) - r[\, A_{21}, \ A_{22}\,] \right\}, \end{aligned}$$

where $G = A_{11}F_{A_{21}}$ *and* $H = E_{A_{21}}A_{22}$.

(b) *The minimal rank of* $M(X_{12})$ *with respect to* X_{12} *is*

$$\min_{X_{12}} r[M(X_{12})] = r\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} + r[\, A_{21}, \ A_{22}\,] - r(A_{21}), \tag{31.7}$$

and the matrix X_{12} *satisfying* (31.7) *is exactly the general solution of the following consistent linear matrix equation*

$$E_G(X_{12} - A_{11}A_{21}^{-}A_{22})F_H = 0,$$

which can be written as

$$X_{12} = A_{11}A_{21}^{-}A_{22} + A_{11}F_{A_{21}}V + WE_{A_{21}}A_{22}, \tag{31.8}$$

where V *and* W *are two arbitrary matrices.*

(c) *The matrix* X_{12} *satisfying* (31.7) *is unique if and only if*

$$r\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} = r[\, A_{21}, \ A_{22}\,] = r(A_{21}).$$

In that case, the unique matrix is $X_{12} = A_{11}A_{21}^{-}A_{22}$.

(d) *The rank of* $M(X_{12})$ *is invariant with respect to the choice of* X_{12}, *if and only if*

$$r(A_{11}) = s_1 \quad \textit{and} \quad R(A_{11}^T) \cap R(A_{21}^T) = \{0\}, \tag{31.9}$$

and

$$r(A_{22}) = t_2 \quad \textit{and} \quad R(A_{21}) \cap R(A_{22}) = \{0\}. \tag{31.10}$$

Proof. Applying (1.6) to $M(X_{12})$ in (31.4), we first obtain

$$r[M(X_{12})] = r\begin{bmatrix} A_{11} & X_{12} \\ A_{21} & A_{22} \end{bmatrix} = r\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} + r[\, A_{21}, \ A_{22}\,] - r(A_{21}) + r[\, E_G(X_{12} - A_{11}A_{21}^{-}A_{22})F_H\,], \tag{31.11}$$

where $G = A_{11}F_{A_{21}}$ and $H = E_{A_{21}}A_{22}$. Thus the maximal and the minimal ranks of $M(X_{12})$ subject to X_{12} are, in fact, determined by the term $E_G(X_{12} - A_{11}A_{21}^{-}A_{22})F_H$. It is quite easy to see that

$$\max_{X_{12}} r[\, E_G(\, X_{12} - A_{11}A_{21}^{-}A_{22}\,)F_H\,] = \min\{\, r(E_G), \quad r(F_H)\,\}, \tag{31.12}$$

and the matrix X_{12} is given by (31.6). Moreover

$$\min_{X_{12}} r[\, E_G(\, X_{12} - A_{11}A_{21}^{-}A_{22}\,)F_H\,] = 0, \tag{31.13}$$

and the matrix X_{12} is given by (31.8). Putting (31.12) and (31.13) in (31.11) produces (31.5) and (31.7). The result in Part (c) is direct consequence of Part (b). The invariance of the rank of $M(X_{12})$ subject to X_{12} is equivalent to the fact that (31.5) and (31.7) are equal, that is,

$$r\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} = r(A_{21}) + s_1 \quad \text{or} \quad r[\, A_{21}, \ A_{22}\,] = r(A_{21}) + t_2.$$

Finally applying Lemma 1.2 to both of them leads to (31.9) and (31.10). □

Notice that $M(X_{12})$ is the simplest case of $A_n + X_n$ in (31.1) corresponding to $n = 2$. Thus Lemma 31.1 presents, in fact, a complete solution to the two problems in (31.2) and (31.3) when $n = 2$. Our work in the next two sections is actually to extend the results in Lemma 1.2 to $n \times n$ case.

31.1. The maximal rank completion of $A_n + X_n$

For convenience of representation, we adopt the following notations for the block matrices in (31.1),

$$P_i = [\, A_{i1},\, A_{i2},\, \cdots,\, A_{i,i-1} \,], \qquad i = 2,\ 3,\ \cdots,\ n, \tag{31.14}$$

$$M_i = [\, A_{i1},\, A_{i2},\, \cdots,\, A_{ii} \,], \qquad i = 1,\, 2,\, \cdots,\, n, \tag{31.15}$$

$$A_i = \begin{bmatrix} A_{11} & & \\ \vdots & \ddots & \\ A_{i1} & \cdots & A_{ii} \end{bmatrix}, \qquad i = 1,\, 2,\, \cdots,\, n, \tag{31.16}$$

$$Q_{ij} = \begin{bmatrix} A_{i1} & \cdots & A_{ij} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nj} \end{bmatrix}, \qquad 1 \le i \le j \le n, \tag{31.17}$$

$$N_i = \begin{bmatrix} A_{ii} \\ A_{i+1,i} \\ \vdots \\ A_{ni} \end{bmatrix}, \qquad Y_i = \begin{bmatrix} X_{1i} \\ X_{2i} \\ \vdots \\ X_{i-1,i} \end{bmatrix}, \quad i = 2,\ 3,\ \cdots,\ n, \tag{31.18}$$

$$X_i = \begin{bmatrix} 0 & X_{12} & \cdots & X_{1i} \\ & 0 & \ddots & \vdots \\ & & \ddots & X_{i-1,i} \\ & & & 0 \end{bmatrix}, \qquad i = 2,\, 3,\, \cdots,\, n-1. \tag{31.19}$$

¿From (31.17) and (31.18) we see that

$$[\, Q_{i+1,i},\ N_{i+1} \,] = Q_{i+1,i+1}, \qquad i = 1,\, 2,\, \cdots,\ n-1. \tag{31.20}$$

Besides, we use

$$S_i = \{\, X_{1,i+1},\, X_{2,i+2},\, \cdots,\, X_{n-i,n} \,\}, \qquad i = 1,\, 2,\, \cdots,\, n-1$$

to denote the set of the $n - i$ variant block entries in the ith upper block subdiagonal of X_n in (31.1).

Theorem 31.2. *Let $A_n + X_n$ be given by* (31.2)*. Then the maximal rank of $A_n + X_n$ subject to $X_n \in S(X_n)$ is*

$$\max_{X_n \in S(X_n)} r(\, A_n + X_n \,)$$
$$= \min \{\, r(Q_{11}) + (s - k_1) + (t - l_1),\ r(Q_{22}) + (s - k_2) + (t - l_2),\ \cdots,\ r(Q_{nn}) + (s - k_n) + (t - l_n) \,\}, \tag{31.21}$$

where s and t are the row number and column number of $A_n + X_n$, respectively, $k_i = \sum_{j=i}^n s_j$ and $l_i = \sum_{j=1}^i t_j$ are the row number and column number of $Q_{ii}(i = 1,\, 2,\, \cdots,\, n)$ in (31.17)*, respectively.*

Proof. By induction on n. When $n = 2$, $A_n + X_n$ in (31.2) has the same form as $M(X_{12})$ in Lemma 31.1, and the result in (31.5) is exactly the result in (31.21) when $n = 2$. Hence (32.21) is true for $n = 2$. Now suppose that (31.21) is true for $A_{n-1} + X_{n-1}$. Then we next consider n. According to (31.14)—(31.19), $A_n + X_n$ in (31.2) can be partitioned as

$$A_n + X_n = \begin{bmatrix} A_{n-1} + X_{n-1} & Y_n \\ Q_{n,n-1} & N_n \end{bmatrix}.$$

In that case, the maximal rank of $A_n + X_n$ subject to $X_n \in S(X_n)$ can be calculated by the following two steps

$$\max_{X_n \in S(X_n)} r(A_n + X_n) = \max_{X_{n-1} \in S(X_{n-1})} \max_{Y_n} r(A_n + X_n). \tag{31.22}$$

applying (31.5), we first find that

$$\max_{Y_n} r(\, A_n + X_n \,) = \max_{Y_n} r \begin{bmatrix} A_{n-1} + X_{n-1} & Y_n \\ Q_{n,n-1} & N_n \end{bmatrix} = \min\left\{ r \begin{bmatrix} A_{n-1} + X_{n-1} \\ Q_{n,n-1} \end{bmatrix} + t_n, \quad r(Q_{nn}) + (s - k_n) \right\}, \tag{31.23}$$

and the matrix Y_n satisfying (31.23) can be written as

$$Y_n = \widehat{Y}_n + (\, A_{n-1} + X_{n-1} \,) Q_{n,n-1}^- N_n + G V_n + W_n H,$$

where V_n and W_n are two arbitrary matrices, $G = (A_{n-1} + X_{n-1}) F_{Q_{n,n-1}}$, $H = E_{Q_{n,n-1}} N_n$, and $\widehat{Y}_n$ is chosen such that

$$r(E_G \widehat{Y}_n F_H) = \min\{\, r(E_G), \quad r(F_H) \,\}.$$

The next step for continuing (31.22) is to find the maximal rank of the block matrix in (31.23) subject to X_{n-1}. Observe that

$$\begin{bmatrix} A_{n-1} + X_{n-1} \\ Q_{n,n-1} \end{bmatrix} = \begin{bmatrix} A_{11} & X_{12} & \cdots & X_{1,n-1} \\ A_{21} & A_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & X_{n-2,n-1} \\ B_1 & B_2 & \cdots & B_{n-1} \end{bmatrix}, \tag{31.24}$$

where $B_i = \begin{bmatrix} A_{n-1,i} \\ A_{ni} \end{bmatrix}$, $i = 1, 2, \cdots, n-1$. Hence (31.24) is, in fact, a new $(n-1) \times (n-1)$ block matrix with the same form as $A_{n-1} + X_{n-1}$ in (31.2). Thus by hypothesis of induction, we know that

$$\max_{X_{n-1} \in S(X_{n-1})} r \begin{bmatrix} A_{n-1} + X_{n-1} \\ Q_{n,n-1} \end{bmatrix}$$
$$= \min\{\, r(Q_{11}) + (\bar{t} - l_1), \quad r(Q_{22}) + (s - k_2) + (\bar{t} - l_2), \quad \cdots, \quad r(Q_{n-1,n-1}) + (s - k_{n-1}) \,\},$$

where $\bar{t} = \sum_{i=1}^{n-1} t_i$. Substituting it into (31.23) yields

$$\max_{X_n \in S(X_n)} r(\, A_n + X_n \,) = \min\{\, r(Q_{11}) + (\bar{t} - l_1) + t_n, \quad r(Q_{22}) + (s - k_2) + (\bar{t} - l_2) + t_n,$$
$$\cdots, \quad r(Q_{n-1,n-1}) + (s - k_{n-1}) + t_n, \quad r(Q_{nn}) + (s - k_n) \,\}.$$

Note that $t = \bar{t} + t_n$, $s - k_1 = 0$ and $t - l_n = 0$, thus the above result is exactly the formula in (31.21). $\square$

¿From the proof of Theorem 31.2 we can also conclude a group of formulas for calculating the column block matrices $Y_2, Y_3, \cdots, Y_n$ in the matrix X_n satisfying (31.21).

Theorem 31.3. *The general expressions of the column block entries Y_2, Y_3, $\cdots$, Y_n in the matrix X_n satisfying* (31.21) *can be written in the inductive formulas*

$$Y_2 = \widehat{Y}_2 + A_{11} Q_{21}^- N_2 + G_2 V_2 + W_2 H_2, \tag{31.25}$$

$$Y_i = \widehat{Y}_i + (\, A_{i-1} + X_{n-1} \,) Q_{i,i-1}^- N_i + G_i V_i + W_i H_i, \quad i = 3, \cdots, n, \tag{31.26}$$

where

$$A_{i-1} + X_{i-1} = \begin{bmatrix} A_{i-2} + X_{i-2} & Y_{i-1} \\ P_{i-1} & A_{i-1,i-1} \end{bmatrix}, \qquad i = 3, \ \cdots, \ n,$$

$V_2, V_3, \cdots, V_n, W_2, W_3, \cdots, W_n$ *are arbitrary matrices,* $G_2 = A_{11}F_{Q_{21}}$, $H_2 = E_{Q_{21}}N_2$ *and*

$$G_i = (A_{i-1} + X_{i-1})F_{Q_{i,i-1}}, \quad H_i = E_{Q_{i,i-1}}N_i, \qquad i = 3, \cdots, n,$$

meanwhile $\widehat{Y}_2, \widehat{Y}_3, \cdots, \widehat{Y}_n$ *are chosen such that*

$$r(E_{G_i}\widehat{Y}_iF_{H_i}) = \min\{\, r(E_{G_i}), \quad r(F_{H_i})\,\}, \qquad i = 2, 3, \cdots, n.$$

Substituting (31.25) and (31.26) into the matrix X_n in (31.2) will produce a general expression for the maximal rank completion of $A_n + X_n$ in (31.2).

On the basis of Theorems 31.3 and 31.3, now we are able to consider the nonsingularity of $A_n + X_n$ in (31.2) when it is a square block matrix.

Corollary 31.4. *Suppose that* $A_n + X_n$ *in* (31.2) *is a square block matrix of size* $t \times t$. *Then there exists an* $X_n \in S(X_n)$ *such that* $A_n + X_n$ *in* (31.2) *is nonsingular, if and only if the block matrices* $Q_{11}, Q_{22}, \cdots, Q_{nn}$ *in* A_n *satisfy*

$$r(Q_{11}) = l_1, \quad r(Q_{22}) \geq k_2 + l_2 - t, \quad \cdots, \quad r(Q_{n-1,n-1}) \geq k_{n-1} + l_{n-1} - t, \quad r(Q_{nn}) = k_n,$$

where k_i *and* l_i *are, respectively, the row number and the column number of* $Q_{ii}(i = 1, 2, \cdots, n)$. *In that case, the column block matrices* $Y_2, Y_3, \cdots, Y_n$ *in the matrix* X_n *such that* $A_n + X_n$ *is nonsingular are also given by* (31.25) *and* (31.26).

If the matrix A_n in (31.1) satisfies some additional conditions, the results in Theorems 31.2 and 31.3 can further simplify. In particular, when A_n in (31.1) is a diagonal block matrix, we have the following simple result.

Corollary 31.5. *Suppose that* A_n *in* (31.1) *is a diagonal block matrix, i.e.,* $A_{ij} = 0(i > j)$ *in* (31.1). *Then*

$$\begin{aligned} &\max_{X_n \in S(X_n)} r(\, A_n + X_n\,) \\ &= \min\{\, r(A_{11}) + (s - k_1) + (t - l_1), \;\; r(A_{22}) + (s - k_2) + (t - l_2), \;\; \cdots, \;\; r(A_{nn}) + (s - k_n) + (t - l_n)\,\}, \end{aligned} \tag{31.27}$$

where k_i *and* $l_i(i = 1, 2, \cdots, n)$ *are as in* (31.21). *The column block entries* $Y_2, Y_3, \cdots, Y_n$ *in the matrix* X_n *satisfying* (31.27) *are given by inductive formulas*

$$\begin{aligned} Y_2 &= \widehat{Y}_2 + A_{11}V_1 + W_1A_{22}, \\ Y_i &= \widehat{Y}_i + (A_{n-1} + X_{n-1})V_i + W_iA_{ii}, \quad i = 3, \cdots, n, \end{aligned}$$

where

$$A_{i-1} + X_{i-1} = \begin{bmatrix} A_{i-2} + X_{i-2} & Y_{i-1} \\ 0 & A_{i-1,i-1} \end{bmatrix}, \quad i = 3, \cdots, n,$$

the matrices $V_2, V_3, \cdots, V_n, W_2, W_3, \cdots, W_n$ *are arbitrary, and* $\widehat{Y}_2, \widehat{Y}_3, \cdots, \widehat{Y}_n$ *in them satisfy*

$$\begin{aligned} &r(E_{A_{11}}\widehat{Y}_2F_{A_{22}}) = \min\{\, r(E_{A_{11}}), \quad r(F_{A_{22}})\,\}, \\ &r(E_{G_i}\widehat{Y}_iF_{A_{ii}}) = \min\{\, r(E_{G_i}), \quad r(F_{A_{ii}})\,\}, \quad i = 3, \cdots, n, \end{aligned}$$

where $G_i = A_{i-1} + X_{i-1}$, $i = 3, \cdots, n$.

31.2. The minimal rank completion of $A_n + X_n$

From the results in Lemma 31.1(b), we see that the variant entry X_{12} such that $M(X_{12})$ has its minimal rank is, in fact, the general solution of a consistent linear matrix equation constructed by the given matrices A_{11}, A_{21} and A_{22} in A_2. It is not difficult to find by repeatedly using Lemma 31.1(b) that the column block entries $Y_2, Y_3, \cdots, Y_n$ in the minimal rank completion of $A_n + X_n$ are also the general solutions of $n - 1$ consistent linear matrix equations constructed by the given block entries in A_n.

Theorem 31.6. *Let* $A_n + X_n$ *be given by* (31.3). *Then*

(a)[152] *The minimal rank of* $A_n + X_n$ *subject to* $X_n \in S(X_n)$ *is*

$$\min_{X_n \in S(X_n)} r(A_n + X_n) = \sum_{i=1}^{n} r(G_{ii}) - \sum_{i=1}^{n-1} r(Q_{i+1,i}). \tag{31.29}$$

(b) *The general expressions of the column block entries* $Y_2, \ Y_3, \ \cdots, \ Y_n$ *in the matrix* X_n *satisfying* (31.3) *can be calculated by the inductive formulas*

$$Y_2 = A_{11}Q_{21}^{-}N_2 + A_{11}F_{Q_{21}}V_2 + W_2E_{Q_{21}}N_2, \tag{31.30}$$

$$Y_i = (A_{i-1} + X_{i-1})Q_{i,i-1}^{-}N_i + G_iV_i + W_iH_i, \quad i = 3, \cdots, n, \tag{31.31}$$

where

$$A_{i-1} + X_{i-1} = \begin{bmatrix} A_{i-2} + X_{i-2} & Y_{i-1} \\ P_{i-1} & A_{i-1,i-1} \end{bmatrix}, \quad i = 3, \ \cdots, \ n,$$

$V_2, V_3, \cdots, V_n, W_2, W_3, \cdots, W_n$ *are arbitrary matrices, and*

$$G_i = (A_{i-1} + X_{i-1})F_{Q_{i,i-1}}, \quad H_i = E_{Q_{i,i-1}}N_i, \quad i = 3, \cdots, n.$$

Proof. According to the structure of X_n in (31.1) we determine the minimal rank completion of $A_n + X_n$ by the following $n-1$ steps

$$\min_{X_n \in S(X_n)} r(A_n + X_n) = \min_{Y_2} \min_{Y_3} \cdots \min_{Y_n} r(A_n + X_n). \tag{31.32}$$

Applying Lemma 31.1(b) we first find that

$$\begin{aligned} \min_{Y_n} r(A_n + X_n) &= \min_{Y_n} r \begin{bmatrix} A_{n-1} + X_{n-1} & Y_n \\ Q_{n,n-1} & N_n \end{bmatrix} \\ &= r(Q_{n,n-1}, \ N_n) - r(Q_{n,n-1}) + r \begin{bmatrix} A_{n-1} + X_{n-1} \\ Q_{n,n-1} \end{bmatrix} \\ &= r(Q_{nn}) - r(Q_{n,n-1}) + r \begin{bmatrix} A_{n-1} + X_{n-1} \\ Q_{n,n-1}, \end{bmatrix}, \end{aligned} \tag{31.33}$$

and the column block matrix Y_n satisfying (31.33) is

$$Y_n = (A_{n-1} + X_{n-1})Q_{n,n-1}^{-}N_n + G_nV_n + W_nH_n, \tag{31.34}$$

where

$$G_n = (A_n + X_n)F_{Q_{n,n-1}}N_n, \quad \text{and} \quad H_n = E_{Q_{n,n-1}}N_n,$$

V_n and W_n are two arbitrary matrices. Clearly (31.34) is exactly the result in (31.31) when $i = n$. Observe that

$$\begin{bmatrix} A_{n-1} + X_{n-1} \\ Q_{n,n-1} \end{bmatrix} = \begin{bmatrix} A_{n-2} + X_{n-2} & Y_{n-1} \\ Q_{n-1,n-2} & N_{n-1} \end{bmatrix}.$$

Thus we find by Lemma 31.1(b) that

$$\min_{Y_{n-1}} r \begin{bmatrix} A_{n-1} + X_{n-1} \\ Q_{n,n-1} \end{bmatrix} = r(Q_{n-1,n-1}) - r(Q_{n-1,n-2}) + r \begin{bmatrix} A_{n-2} + X_{n-2} \\ Q_{n-1,n-2} \end{bmatrix}, \tag{31.35}$$

and Y_{n-1} satisfying (31.35) is

$$Y_n = (A_{n-2} + X_{n-2})Q_{n-1,n-2}^{-}N_{n-1} + G_{n-1}V_{n-1} + W_{n-1}H_{n-1}, \tag{31.36}$$

where

$$G_{n-1} = (A_{n-2} + X_{n-2})F_{Q_{n-1,n-2}}, \quad \text{and} \quad H_{n-1} = E_{Q_{n-1,n-2}}N_{n-1}, \ V_{n-1},$$

and W_{n-1} are two arbitrary matrices. Clearly (31.36) is exactly the result in (31.31) when $i = n-1$. By the same method, we can derive a group of inductive formulas as follows

$$\min_{Y_i} r \begin{bmatrix} A_i + X_i \\ Q_{i+1,i} \end{bmatrix} = r(Q_{ii}) - r(Q_{i,i-1}) + r \begin{bmatrix} A_{i-1} + X_{i-1} \\ Q_{i,i-1} \end{bmatrix}, \qquad i = n-1, \cdots, 3,$$

and

$$\min_{Y_2} r \begin{bmatrix} A_2 + X_2 \\ Q_{32} \end{bmatrix} = r(Q_{11}) + r(Q_{22}) - r(Q_{21}).$$

The general expressions of $Y_2, Y_3, \cdots, Y_n$ satisfying the above group of equalities are given in (31.30) and (31.31). Finally substituting the above $n-1$ equalities into (31.32) results in (31.29). □

A particular concern for the problem (31.3) is the uniqueness of X_n satisfying (31.3), which was well examined in [47], [152] and [153].

Corollary 31.7. *Let $A_n + X_n$ be given by* (31.2).

(a)[152] *The matrix X_n satisfying* (31.1) *is unique, that is, the matrix $A_n + X_n$ has a unique minimal rank completion, if and only if*

$$r(Q_{ii}) = r(Q_{i+1,i}) = r(Q_{i+1,i+1}), \qquad i = 1, 2, \cdots, n-1, \tag{31.37}$$

In that case, the minimal rank of $A_n + X_n$ subject to $X_n \in S(X_n)$ is

$$\min_{X_n \in S(X_n)} r(\, A_n + X_n \,) = r(Q_{11}). \tag{31.38}$$

(b) *Under the conditions in* (31.37), *the column block entries $Y_2,\ Y_3, \cdots, Y_n$ in the unique minimal rank completion of $A_n + X_n$ are given by the group of inductive formulas*

$$Y_2 = A_{11} Q_{21}^{-} N_2, \qquad Y_i = (\, A_{i-1} + X_{i-1} \,) Q_{i,i-1}^{-} N_i, \qquad i = 3, \cdots, n,$$

where

$$A_{i-1} + X_{i-1} = \begin{bmatrix} A_{i-2} + X_{i-2} & Y_{i-1} \\ P_{i-1} & A_{i-1,i-1} \end{bmatrix}, \qquad i = 3, \cdots, n.$$

Proof. Follows directly from (31.30) and (31.31). □

In general cases, the maximal rank and the minimal rank of $A_n + X_n$ subject to $X_n \in S(X_n)$ are different. If they are equal, it implies that the rank of $A_n + X_n$ is invariant for any choice of $X_n \in S(X_n)$. The combination of Theorem 31.2 with Theorem 31.6 yields the following result.

Corollary 31.8. *Let $A_n + X_n$ be given by* (31.2). *Then the rank of $A_n + X_n$ is invariant with respect to the choice of $X_n \in S(X_n)$ if and only if*

$$\begin{aligned} &\sum_{i=1}^{n} r(Q_{ii}) - \sum_{i=1}^{n-1} r(Q_{i+1,i}) \\ &= \min\{ r(Q_{11}) + (s-k_1) + (t-l_1),\ r(Q_{22}) + (s-k_2) + (t-l_2),\ \cdots,\ r(Q_{nn}) + (s-k_n) + (t-l_n) \}, \end{aligned} \tag{31.39}$$

where s and t are the row and the column numbers of $A_n + X_n$, respectively; k_i and l_i are the row and the column numbers of $Q_{ii} (i = 1, 2, \cdots, n)$, respectively. In that case, any sum of of $A_n + X_n$ can be regarded as a maximal or a minimal rank completion of the triangular block matrix A_n.

The rank equality (31.39) can be written as some explicit equivalent expressions. When $n = 2$, the rank equality in (31.39) is Lemma 31.1(d). We next present a equivalent statement for (31.39) when $n = 3$. The proof is analogous to that of Lemma 31.1(d) and is, therefore, omitted.

Corollary 31.9. *Suppose that $A_3 + X_3$ is a 3×3 block matrix as follows*

$$A_3 + X_3 = \begin{bmatrix} A_{11} & X_{12} & X_{13} \\ A_{21} & A_{22} & X_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \tag{31.40}$$

Then the rank of $A_3 + X_3$ *is invariant with respect to the choice of* X_{12}, X_{13} *and* X_{23}, *if and only if one of the following three groups of conditions is satisfied:*

(a) A_{11} *and* $[\, A_{21},\ A_{22}\,]$ *have full row ranks, respectively, and*

$$R(A_{11}^T) \cap R[\, A_{21}^T,\ A_{31}^T\,] = \{0\}, \quad \textit{and} \quad R\begin{bmatrix} A_{21}^T \\ A_{22}^T \end{bmatrix} \cap R\begin{bmatrix} A_{31}^T \\ A_{32}^T \end{bmatrix} = \{0\}.$$

(b) A_{11} *has full row rank,* A_{33} *has full column rank, and*

$$R(A_{11}^T) \cap R[\, A_{21}^T,\ A_{31}^T\,] = \{0\}, \quad \textit{and} \quad R[\, A_{31},\ A_{32}\,] \cap R(A_{33}) = \{0\}.$$

(c) A_{33} *and* $\begin{bmatrix} A_{22} \\ A_{32} \end{bmatrix}$ *has full column ranks, respectively, and*

$$R\begin{bmatrix} A_{21} \\ A_{31} \end{bmatrix} \cap R\begin{bmatrix} A_{22} \\ A_{32} \end{bmatrix} = \{0\}, \quad \textit{and} \quad R[\, A_{31},\ A_{32}\,] \cap R(A_{33}) = \{0\}.$$

Remark 31.10. Besides the partitioning method for X_n shown in (31.18) and (31.19), we can also inductively partition X_n in (31.2) into

$$X_n = \begin{bmatrix} 0 & Z_1 \\ 0 & X_{n-1} \end{bmatrix}, \qquad \cdots, \qquad X_3 = \begin{bmatrix} 0 & Z_{n-2} \\ 0 & X_2 \end{bmatrix}, \qquad X_2 = \begin{bmatrix} 0 & X_{n-1,n} \\ 0 & 0 \end{bmatrix},$$

where $Z_i = [\, X_{i,i+1},\, X_{i,i+2},\, \cdots,\, X_{in}\,]$, $i = 1,\, 2,\, \cdots,\, n-1$. In the same method for deducing the results in Sections 31.1 and 31.2, we can also find general expressions of the row block matrices $Z_1,\ Z_2,\ \cdots,\ Z_{n-1}$ to the two problems in (31.2) and (31.3). The corresponding results are much analogous to those in Theorems 31.2 and 31.3, and are, therefore, omitted here.

In addition to the two methods used or mentioned above, another method available for constructing maximal and minimal rank completions of $A_n + X_n$ in (31.2) and (31.3) is inductively calculating the block entries in each upper block subdiagonal of X_n. We next illustrate this method by constructing maximal and minimal rank completions of $A_3 + X_3$ in (31.40).

According to Lemma 31.1(a) and (b), the maximal and the minimal ranks of $A_3 + X_3$ in (31.40) with respect to X_{13} respectively are

$$\max_{X_{13}} r(\, A_3 + X_3\,) = \min\left\{ r\begin{bmatrix} A_{11} & X_{12} \\ Q_{21} & N_2 \end{bmatrix} + t_3, \quad r\begin{bmatrix} M_2 & X_{23} \\ Q_{32} & A_{33} \end{bmatrix} + s_1 \right\}, \tag{31.41}$$

and

$$\min_{X_{13}} r(\, A_3 + X_3\,) = r\begin{bmatrix} A_{11} & X_{12} \\ Q_{21} & N_2 \end{bmatrix} + r\begin{bmatrix} M_2 & X_{23} \\ Q_{32} & A_{33} \end{bmatrix} - r(Q_{22}). \tag{31.42}$$

The general expressions of X_{13} satisfying (31.41) and (31.42) can respectively be derived from (31.6) and (31.8). Observe that the variant entries X_{12} and X_{23} occur in two independent block matrices in (31.41) and (31.42). Applying Lemma 31.1(a) and (b) to the block matrices in (31.41) and (31.42), we easily obtain

$$\begin{aligned} \max_{X_3 \in S(X_3)} r(A_3 + X_3) &= \min\left\{ \max_{X_{12}} r\begin{bmatrix} A_{11} & X_{12} \\ Q_{21} & N_2 \end{bmatrix} + t_3, \quad \max_{X_{23}} r\begin{bmatrix} M_2 & X_{23} \\ Q_{32} & A_{33} \end{bmatrix} + s_1 \right\} \\ &= \min\{ r(Q_{11}) + t_2 + t_3, \quad r(Q_{22}) + s_1 + t_3, \quad r(Q_{33}) + s_1 + s_2 \}, \end{aligned}$$

and

$$\begin{aligned} \min_{X_3 \in S(X_3)} r(A_3 + X_3) &= \min_{X_{21}} r\begin{bmatrix} A_{11} & X_{12} \\ Q_{21} & N_2 \end{bmatrix} + \min_{X_{23}} r\begin{bmatrix} M_2 & X_{23} \\ Q_{32} & A_{33} \end{bmatrix} - r(Q_{22}) \\ &= r(Q_{11}) + r(Q_{22}) + r(Q_{33}) - r(Q_{21}) - r(Q_{32}). \end{aligned}$$

The matrices X_{12} and X_{23} satisfying the above two equalities can respectively be derived from (31.6) and (31.7). Clearly the above two equalities are exactly (31.27) and (31.29) when $n = 3$ in them.

It is easy to conclude from the above example that the two completion problems can also be constructed by inductively calculating the block entries in the upper block subdiagonals in X_n. Speaking precisely, the first step of this work is to find the $n-1$ block entries in the set $S_1 = \{ X_{21}, X_{23}, \cdots, X_{n-1,n} \}$ of X_n such that

$$r\begin{bmatrix} M_1 & X_{12} \\ Q_{21} & N_2 \end{bmatrix} = \max, \qquad r\begin{bmatrix} M_2 & X_{23} \\ Q_{32} & N_3 \end{bmatrix} = \max, \quad \cdots, \quad r\begin{bmatrix} M_{n-1} & X_{n-1,n} \\ Q_{n,n-1} & N_n \end{bmatrix} = \max,$$

and

$$r\begin{bmatrix} M_1 & X_{12} \\ Q_{21} & N_2 \end{bmatrix} = \min, \qquad r\begin{bmatrix} M_2 & X_{23} \\ Q_{32} & N_3 \end{bmatrix} = \min, \quad \cdots, \quad r\begin{bmatrix} M_{n-1} & X_{n-1,n} \\ Q_{n,n-1} & N_n \end{bmatrix} = \min$$

hold, respectively, where $M_i,\ Q_{i,i-1}$ and N_i are defined in (31.14)—(31.17). The second step is to substitute the $n-1$ given block entries in S_1 into $A_n + X_n$ and then to determine the $n-2$ block entries in the set $S_2 = \{X_{13},\ X_{24},\ \cdots,\ X_{n-2,n}\}$ of X_n such that

$$r\begin{bmatrix} * & X_{13} \\ * & * \end{bmatrix} = \max, \qquad r\begin{bmatrix} * & X_{24} \\ * & * \end{bmatrix} = \max, \qquad \cdots, \qquad r\begin{bmatrix} * & X_{n-2,n} \\ * & * \end{bmatrix} = \max$$

and

$$r\begin{bmatrix} * & X_{13} \\ * & * \end{bmatrix} = \min, \qquad r\begin{bmatrix} * & X_{24} \\ * & * \end{bmatrix} = \min, \qquad \cdots, \qquad r\begin{bmatrix} * & X_{n-2,n} \\ * & * \end{bmatrix} = \min$$

hold, respectively, where $\begin{bmatrix} * & X_{ij} \\ * & * \end{bmatrix}$ denotes the submatrix in $A_n + X_n$ as follows

$$\begin{bmatrix} * & X_{ij} \\ * & * \end{bmatrix} = \begin{bmatrix} A_{i1} & \cdots & X_{ij} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nj} \end{bmatrix}, \quad 1 \le i < j \le n.$$

Next substituting the $n-2$ given block matrices in S_2 into A_n+X_n and repeating the analogous calculations produces inductively the block entries in the 3rd, 4th, ..., $(n-1)$th upper block subdiagonals in X_n. Finally substituting all the given entries in the $n-1$ upper block subdiagonals of X_n into $A_n + X_n$ yields maximal and minimal rank completions of $A_n + X_n$, respectively.

Summing up all the results and discussion given in Sections 31.1 and 31.2, we see that there are three kinds of general methods available for constructing maximal and minimal rank completions of a triangular block matrix:

(i) Calculating inductively the unspecified column block entries $Y_2, Y_3, \cdots.\ Y_n$ in $A_n + X_n$.

(ii) Calculating inductively the unspecified row block entries $Z_1, Z_2, \cdots,\ Z_{n-1}$ in $A_n + X_n$.

(iii) Calculating inductively the unspecified upper block subdiagonal entries in $A_n + X_n$.

31.3. Extreme ranks of $A - BXC$ when X is a triangular block matrix

Let $A - BXC$ be a matrix expression over an arbitrary field $\mathcal{F}$, and suppose X is a variant $p \times p$ upper triangular block matrix. In that case, $A - BXC$ can also be written as

$$A - BXC = A - [\,B_1,\, B_2,\, \cdots,\, B_p\,] \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ & X_{22} & \cdots & X_{2p} \\ & & \ddots & \vdots \\ & & & X_{pp} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_p \end{bmatrix}, \tag{31.43}$$

where $A \in \mathcal{F}^{m\times n}$, $B \in \mathcal{F}^{m\times k}$, $C \in \mathcal{F}^{l\times n}$, $B_i \in \mathcal{F}^{m\times k_i}$, $C_i \in \mathcal{F}^{l_i\times n}$, $X_{ij} \in \mathcal{F}^{k_i\times l_j}(1 \le i \le j \le p)$.

In this section, we consider how to determine the maximal and the minimal ranks of the matrix expression (31.43) with respect to all variant blocks X_{ij}. This work is motivated by some earlier work on

triangular block solutions of matrix equations, and triangular block generalized inverses of matrices (see, e.g., [91], [92]).

It is easy to verify that the rank of $A - BXC$ can be expressed as the rank of a block matrix as follows

$$r(A - BXC) = r\begin{bmatrix} 0 & I_k & X \\ C & 0 & I_l \\ -A & B & 0 \end{bmatrix} - k - l.$$

Putting (31.43) in it, we get

$$\begin{aligned} r(A - BXC) &= r\begin{bmatrix} 0 & I_{k_1} & 0 & \cdots & 0 & X_{11} & X_{12} & \cdots & X_{1p} \\ 0 & 0 & I_{k_2} & \cdots & 0 & 0 & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_{k_p} & 0 & 0 & \cdots & X_{pp} \\ C_1 & 0 & 0 & \cdots & 0 & I_{l_1} & 0 & \cdots & 0 \\ C_2 & 0 & 0 & \cdots & 0 & 0 & I_{l_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ C_p & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & I_{l_p} \\ -A & B_1 & B_2 & \cdots & B_p & 0 & 0 & \cdots & 0 \end{bmatrix} - k - l \\ &= r\begin{bmatrix} G_{11} & X_{11} & X_{12} & \cdots & X_{1p} \\ G_{21} & G_{22} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ G_{p1} & G_{p2} & G_{p3} & \cdots & X_{pp} \\ G_{p+1,1} & G_{p+1,2} & G_{p+1,3} & \cdots & G_{p+1,p+1} \end{bmatrix} - k - l. \end{aligned} \tag{31.44}$$

This equality shows that the maximal and the minimal ranks of $A - BXC$ in (31.43) can completely determined by those of the block matrix in (31.43). Applying (31.27) and (31.29) to the block matrix in (31.44) and then simplifying, we obtain the following.

Theorem 31.11. *Let $A - BXC$ be given by* (31.43), *and let*

$$\widehat{B}_i = [\, B_1, \cdots, B_i \,], \qquad \widehat{C}_i = \begin{bmatrix} C_i \\ \vdots \\ C_p \end{bmatrix}, \qquad i = 1, 2 \cdots, p.$$

Then

$$\max_{X_{ij}} r(A - BXC) = \min\left\{ r\begin{bmatrix} A \\ \widehat{C}_1 \end{bmatrix}, \; r\begin{bmatrix} A & \widehat{B}_1 \\ \widehat{C}_2 & 0 \end{bmatrix}, \; \cdots, \; r\begin{bmatrix} A & \widehat{B}_{p-1} \\ \widehat{C}_p & 0 \end{bmatrix}, \; r[\, A, \; \widehat{B}_p \,] \right\}, \tag{31.45}$$

and

$$\begin{aligned} &\min_{X_{ij}} r(A - BXC) \\ &= r\begin{bmatrix} A \\ \widehat{C}_1 \end{bmatrix} + r\begin{bmatrix} A & \widehat{B}_1 \\ \widehat{C}_2 & 0 \end{bmatrix} + \cdots + r\begin{bmatrix} A & \widehat{B}_{p-1} \\ \widehat{C}_p & 0 \end{bmatrix} + r[\, A, \; \widehat{B}_p \,] - r\begin{bmatrix} A & \widehat{B}_1 \\ \widehat{C}_1 & 0 \end{bmatrix} - \cdots - r\begin{bmatrix} A & \widehat{B}_p \\ \widehat{C}_p & 0 \end{bmatrix}. \end{aligned} \tag{31.46}$$

In particular, the upper triangular block matrix X satisfying (31.45) *is unique if and only if*

$$r(G_{ii}) = r(G_{i+1,i}) = r(G_{i+1,i+1}), \qquad i = 1, 2, \cdots, p, \tag{31.47}$$

where G_{ij} is defined in (31.44).

For simplicity, the steps for presenting (31.45) and (31.46) are omitted. Furthermore the upper triangular block matrix X satisfying (31.45) and (31.46) can respectively be determined by the three general methods presented in Sections 31.1 and 2. We also omit them here for simplicity.

Two direct consequences of the formula (31.46) are given below.

Corollary 31.12. *The matrix equation*

$$[\,B_1,\, B_2,\, \cdots,\, B_p\,] \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ & X_{22} & \cdots & X_{2p} \\ & & \ddots & \vdots \\ & & & X_{pp} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_p \end{bmatrix} = A \tag{31.48}$$

is consistent if and only if $R(A) \subseteq R(B)$, $R(A^T) \subseteq R(C^T)$, *and*

$$r\begin{bmatrix} A & B_1 & \cdots & B_i \\ C_{i+1} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ C_p & 0 & \cdots & 0 \end{bmatrix} = r\begin{bmatrix} C_{i+1} \\ \vdots \\ C_p \end{bmatrix} + r[\,B_1,\ \cdots,\ B_i\,], \qquad i = 1, 2 \cdots, p-1. \tag{31.49}$$

Proof. Let the right-hand side of (31.46) be zero and then simplify to yield the desired result. □

The general expressions of X_{ij} satisfying (31.48) can also be determined by the general method presented in Sections 31.2.

It is well known that an inner inverse of a matrix A is a solution to the matrix equation $AXA = A$. Thus applying Corollary 31.12 to the equation $AXA = A$, we obtain the following.

Corollary 31.13. *Let*

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ A_{21} & A_{22} & \cdots & A_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pp} \end{bmatrix}. \tag{31.50}$$

Then A has an inner inverse with the upper triangular block form

$$A^- = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ & S_{22} & \cdots & S_{2p} \\ & & \ddots & \vdots \\ & & & S_{pp} \end{bmatrix}, \tag{31.51}$$

if and only if

$$r(A) = r(Q_{1k}) + r(Q_{k+1,p}) - r(Q_{k+1,k}), \qquad k = 1, 2, \cdots, p-1, \tag{31.52}$$

where

$$Q_{ij} = \begin{bmatrix} A_{i1} & \cdots & A_{ij} \\ \vdots & \ddots & \vdots \\ A_{p1} & \cdots & A_{pj} \end{bmatrix} \in \mathcal{F}^{s_i \times t_j}, \qquad 1 \le i,\ j \le p.$$

In particular, an upper triangular block matrix

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ & A_{22} & \cdots & A_{2p} \\ & & \ddots & \vdots \\ & & & A_{pp} \end{bmatrix}$$

has an upper triangular block inner inverse of the form (31.51), *if and only if*

$$r(A) = r(\widehat{A}_{ii}) + r(\tilde{A}_{ii}), \quad i = 1, 2, \cdots, p-1.$$

where

$$\widehat{A}_{ii} = \begin{bmatrix} A_{11} & \cdots & A_{1i} \\ & \ddots & \vdots \\ & & A_{ii} \end{bmatrix}, \quad \textit{and} \quad \tilde{A}_{ii} = \begin{bmatrix} A_{i+1,i+1} & \cdots & A_{i+1,p} \\ & \ddots & \vdots \\ & & A_{pp} \end{bmatrix}.$$

Department of Mathematics and Statistics
Queen's University
Kingston, Ontario, Canada K7L 3N6
e-mail address: ytian@mast.queensu.ca

Bibliography

[1] T. W. Anderson and G. P. H. Styan, Cochran's theorem, rank additivity and tripotent matrices. In Statistics and Probability: Essays in Honor of C. R. Rao (G. Kallianpur et al, eds.), North-Holland, Amsterdam, 1982, pp. 1–23.

[2] W. N. Anderson, Shorted operators, *SIAM J. Appl. Math.* 20(1971), 520-525.

[3] T. Ando, Generalized Schur complements, Linear Algebra Appl. 27(1979), 173-186.

[4] E. Arghiriade, Remarques sur l'inverse généralisée d'un produit de matrices, *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* 42(1967), 621-625.

[5] J. K. Baksalary and P. Kala, The matrix equation $AXB+CYD=E$, *Linear Algebra Appl.* 30(1980), 141-147.

[6] J. K. Baksalary and R. Kala, Range invariance of certain matrix products, *Linear and Multilinear Algebra* 14(1986), 89-96.

[7] J. K. Baksalary and T. Mathew, Rank invariance criterion and its application to the unified theory of least squares, *Linear Algebra Appl.* 127(1990), 393-401.

[8] J. K. Baksalary, F. Pukelsheim and G. P. H. Styan, Some properties of matrix partial orderings, *Linear Algebra Appl.* 119(1989), 57-85.

[9] J. Z. Baksalary and G. P. H. Styan, Around a formula for the rank of a matrix product with some statistical applications. Graphs, matrices, and designs, Lecture Notes in Pure and Appl. Math., 139, Dekker, New York, 1993, pp. 1-18.

[10] C. S. Ballantine, Products of EP matrices, *Linear Algebra and Appl.* 12(1975), 257-267.

[11] W. Barrett, M. Lundquist, C. R. Johnson and H. J. Woerdeman, Completing a block diagonal matrix with partially prescribed inverse, *Linear Algebra Appl.* 223/224(1995), 73-87.

[12] D. T. Barwick and J. D. Gilbert, Generalization of the reverse order law with related results, *Linear Algebra Appl.* 8(1974), 345-349.

[13] D. T. Barwick and J. D. Gilbert, Generalization of the reverse order law with related results, *Linear Algebra Appl.* 8(1974), 345–349.

[14] D. T. Barwick and J. D. Gilbert, On generalizations of the reverse order law with related results, *SIAM. J. Appl. Math.* 27(1974), 326-330.

[15] C. L. Bell, Generalized inverses of circulant and generalized circulant matrices, *Linear Algebra Appl.* 39(1981), 133-142.

[16] A. Ben-Israel and T. N. E. Greville, Generalized inverses: theory and applications. Corrected reprint of the 1974 original. Robert E. Krieger Publishing Co., Inc., Huntington, New York, 1980.

[17] J. Bérubé, R. E. Hartwig and G. P. H. Styan, On canonical correlations and the degrees of non-orthogonality in the three-way layout. In Statistical Sciences and Data Analysis: Proceedings of the Third Pacific Area Statistical Conference, Tokyo, 1991 (K. Matusita et al., eds.), VSP International Science Publishers, Utrecht, The Netherlands, 1993, pp. 247–252.

[18] P. Bhimasankaram, On generalized inverses of partitioned matrices, *Sankhyā Ser. A* 33(1971), 311-314.

[19] F. Burns, D. Carlson, E. Haynsworth and T. Markham, Generalized inverse formulas using the Schur complements, *SIAM. J. Appl. Math.* 26(1974), 254-259.

[20] S. L. Campbell and C. D. Meyer, Jr., EP operators and generalized inverses, *Canad. Math. Bull.* 18(1975), 327-333.

[21] S. L. Campbell and C. D. Meyer, Jr., Generalized inverses of linear transformations, Corrected reprint of the 1979 original. Dover Publications, Inc., New York, 1991.

[22] S. L. Carcdus, Generalized Inverses and Operator Theory, Queen's Papers in Pure and Applied Mathematics No. 50(1978).

[23] D. Carlson, What are Schur complements, anyway? *Linear Algebra Appl.* 74(1986), 257-275.

[24] D. Carlson, E. Haynsworth and T. Markham, A generalization of the Schur complement by means of the Moore-Penrose inverse, *SIAM J. Appl. Math.* 26(1974), 169-175.

[25] X. Chen and R. E. Hartwig, The group inverse of a triangular matrix, *Linear Algebra Appl.* 237/238(1996), 97-108.

[26] Y. Chen, The generalized Bott-Duffin inverse and its applications, *Linear Algebra Appl.* 134(1990), 71-91.

[27] Y. Chen and B. Zhou, On the g-inverses and nonsingularity of a bordered matrix $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$, *Linear Algebra Appl.* 133(1990), 133-151.

[28] L. Cheng and J. B. Pearson, Jr, Synthesis of linear multivariable regulators, *IEEE Trans. Automat. Control* 26(1981), 194-202.

[29] D. Chillag, Generalized circulants and class functions of finite groups, Linear Algebra Appl. 93 (1987), 191-208.

[30] R. E. Cline, Note on the generalized inverse of the product of matrices, *SIAM Rev.* 6(1964), 57-58.

[31] R. E. Cline, Representations of the generalized inverse of sums of matrices,*SIAM J. Numer. Anal. Ser. B* 2(1965), 99-114.

[32] R. E. Cline, Inverse of rank invariant powers of a matrix, *SIAM J. Numer. Anal.* 5(1968), 182-197.

[33] R. E. Cline and R. E. Funderlic, A theorem on the rank of a difference of matrices, *Bull. Amer. Math. Soc.* 82(1976), 48.

[34] R. E. Cline and R. E. Funderlic, The rank of a difference of matrices and associated generalized inverses, *Linear Algebra Appl.* 24(1979), 185-215.

[35] R. E. Cline and T. N. E. Greville, An extension of the generalized inverse of a matrix, *SIAM J. Appl. Math.* 19(1972), 682-688.

[36] R. E. Cline, R. J. Plemmons and G. Worm, generalized inverse of certain Toeplitz matrices, *Linear Algebra Appl.* 8(1974), 25-33.

[37] N. Cohen, C. R. Johnson, L. Rodman and H. J. Woerdeman, Ranks of completions of partial matrices, *Oper. Theory: Adv. Appl.* 40(1989), 165-185.

[38] C. Davis, Completing a matrix so as to minimize its rank, *Operator Theory: Advances and Appl.* 29(1988), 87-95.

[39] P. J. Davis, Circulant Matrices, Wiley, New York, 1979.

[40] T. De Mazancourt, D. Gerlic, The inverse of a block-circulant matrix, *IEEE Trans. Antennas and Propagation* 31(1983), 808-810.

[41] A. R. De Pierro and M. Wei, Reverse order law for reflexive generalized inverses of products of matrices, *Linear Algebra Appl.* 277(1998), 299–311.

[42] L. Elsner and Kh. D. Ikramov, Normal matrices: an update, *Linear Algebra and Appl.* 285(1998), 291-303.

[43] I. Erdelyi, On the "reverse order law" related to the generalized inverse of matrix products, *J. Assoc. Comput. Mach.* 13(1966), 439-433.

[44] M. Fiedler, Remarks on the Schur complement, *Linear Algebra Appl.* 39(1981), 189-195.

[45] M. Fiedler and T. L. Markham, Completing a matrix when certain entries of its inverse are specified, *Linear Algebra Appl.* 74(1986), 225-237.

[46] A. J. Getson and F. C. Hsuan, {2}-Inverses of Matrices and Their Statistical Applications, Lecture Notes in Statistics 47, Springer, Berlin, 1988.

[47] I. Gohberg, M. Kaashoek and L. Lerer, On minimality in the partial realization problem, *Systems Control Lett.* 9(1987), 97-104.

[48] H. Goller, Shorted operators and rank decomposition matrices, *Linear Algebra Appl.* 81(1986), 207-236.

[49] M. C. Gouvreia and R. Puystiens, About the group inverse and Moore-Penrose inverse of a product, *Linear Algebra Appl.* 150(1991), pp. 361-369.

[50] T. N. E. Greville, Note on the generalized inverse of a matrix product, *SIAM Rev.* 8(1966), 518-521.

[51] T. N. E. Greville, Solutions of the matrix equation $XAX = X$, *SIAM J. Appl. Math.* 26(1974), 828-832.

[52] R. Grone, C. R. Johnson, E. D. Sa and H. Wolkowicz, Normal matrices, *Linear Algebra and Appl.* 285(1998), 291-303.

[53] J. Gross, Comments on range invariance of matrix products, *Linear and Multilinear Algebra* 41(1996), 157-160.

[54] J. Gross, Some remarks concerning the reverse order law, *Discuss. Math. Algebra Stochastic Methods* 17(1997), 135-141.

[55] J. Gross and G. Trenkler, Generalized and hypergeneralized projectors, *Linear Algebra Appl.* 87(1987), 213-215.

[56] R. E. Hartwig, The resultant and the matrix equation $AX = XB$, *SIAM J. Appl. Math.* 22(1972), 538-544.

[57] R. E. Hartwig, Block generalized inverses,*Arch. Rational Mech. Anal.* 61(1976), 197-251.

[58] R. E. Hartwig, Singular value decompositions and the Moore-Penrose inverse of bordered matrices, *SIAM J. Appl. Math.* 31(1976), 31–41.

[59] R. E. Hartwig, Rank factorization and the Moore-Penrose inversion, *J. Indust. Math. Soc.* 26(1976), 49-63.

[60] R. E. Hartwig, The reverse order law revisited, Linear Algebra Appl. 76(1986), 241-246.

[61] R. E. Hartwig and I. J. Katz, On products of EP matrices, *Linear Algebra Appl.* 252(1997), 338-345.

[62] R. E. Hartwig, M. Omladič, P. Šemrl and G. P. H. Styan, On some characterizations of pairwise star orthogonality using rank and dagger additivity and subtractivity. Special issue honoring C. R. Rao, *Linear Algebra Appl.* 237/238(1996), 499-507.

[63] R. E. Hartwig and K. Spindelböck, Partial isometries, contractions and EP matrices, *Linear and Multilinear Algebra* 13(1983), 295–310.

[64] R. E. Hartwig and K. Spindelböck, Matrices for which A^* and $A^\dagger$ can commute, *Linear and Multilinear Algebra* 14(1984), 241-256.

[65] R. E. Hartwig and G. P. H. Styan, On some characterizations of the "star" partial ordering for matrices and rank subtractivity, *Linear Algebra Appl.* 82(1986), 145-161.

[66] R. E. Hartwig and G. P. H. Styan, Partially ordered idempotent matrices. In Proceedings of the Second International Tampere Conference in Statistics (T. Pukkila and S. Puntanen, eds.), 1987, pp. 361-383.

[67] E. V. Haynsworth, Applications of an inequality for the Schur complement, *Proc. Amer. Math. Soc.* 24(1970), 512-516.

[68] D. A. Harville, Generalized inverses and ranks of modified matrices, *J. Indian Soc. Agricultural Statist.* 49(1996/1997), 67-78.

[69] C. He, The general solution of the matrix equation $AXB + CYD = F$, *Acta. Sci. Natur. Univ. Norm. Hunan* 19(1996), 17-20.

[70] R. A. Horn and C. R. Johnson, Topics in matrix analysis, Cambridge University Press, Cambridge, 1991.

[71] F. C. Hsuan, P. Langenberg and A. J. Getson, The $\{2\}$ inverse with applications to statistics,*Linear Algebra Appl.* 70(1985), 241-248.

[72] D. Hu, The general solution to the matrix equation $AXB + CYD = E$, *Math. Practice Theory* no. 4(1992), 85-87.

[73] C. H. Hung and T. L. Markham, The Moore-Penrose inverse of a partitioned matrix $\begin{bmatrix} A & C \\ B & D \end{bmatrix}$, *Linear Algebra Appl.* 11(1975), 73-86.

[74] C. H. Hung and T. L. Markham, The Moore-Penrose inverse of a sum of matrices, *J. Austral. Math. Soc. Ser. A* 24(1977), 385-392.

[75] C. R. Johnson, Matrix Completion Problems: A survey, In Matrix Theory and Applications, *Proc. Sympos. Appl. math. AMS* 0(1990), 171-197.

[76] C. R. Johnson and G. T. Whitney, Minimum rank completions, *Linear and Multilinear Algebra* 28(1991), 271-273.

[77] M. Kaashoek and H. J. Woerdeman, Unique minimal rank extension of triangular operators, *J. Math. Anal. Appl.* 131(1988), 501-516.

[78] I. J. Katz, Wiegmann type theorems for EP_r matrices, *Duke Math. J.* 32(1965), 423-427.

[79] I. J. Katz and M. H. Pearl, On EP_r matrices and normal EP_r matrices, *J. Res. Nat. Bur. Standards, Sec B.* 70B(1966), 47-77.

[80] E. P. Liski and S. Wang, On the $\{2\}$-inverse and some ordering properties of nonnegative definite matrices, *Acta Math. Appl. Sinica(English Ser.)* 12(1996), 8-13.

[81] G. Marsaglia and G. P. H. Styan, When does $\mathrm{rank}(A+B) = \mathrm{rank}(A)+\mathrm{rank}(B)$? *Canad. Math. Bull.* 15(1972), 451-452.

[82] G. Marsaglia and G. P. H. Styan, Equalities and inequalities for ranks of matrices, *Linear and Multilinear Algebra* 2(1974), 269-292.

[83] G. Marsaglia and G. P. H. Styan, Rank conditions for generalized inverses of partitioned matrices, *Sankhyā Ser. A 36* (1974), 437-442.

[84] A. R. Meenakshi and R. Indira, On sums of conjugate EP matrices, *Indian J. Pure Appl. Math.* 23(1992), 179-184.

[85] A. R. Meenakshi and R. Indira, Conjugate EP_r factorization of a matrix, *Math. Student* 61(1992), 136-144.

[86] A. R. Meenakshi and R. Indira, On products of conjugate EP_r matrices, *Kyungpook Math. J.* 32(1992), 103-110.

[87] A. R. Meenakshi and R. Indira, On Schur complements in a conjugate EP matrix,*Studia Sci. Math. Hungar.* 32(1996),31-39.

[88] A. R. Meenakshi and R. Indira, On conjugate EP matrices, *Kyungpook Math. J.* 37(1997), 67-72.

[89] A. R. Meenakshi and S. Krishnamoorthy, On k-EP matrices, *Linear Algebra Appl.* 269(1998), 219-232.

[90] A. R. Meenakshi and C. Rajian, On sums and products of star-dagger matrices, *J. Indian Math. Soc.* 50(1988), 149-156.

[91] C. D. Meyer, Jr., Generalized inverses of triangular matrices, *SIAM J. Appl. Math.* 18(1970), 401-406.

[92] C. D. Meyer, Jr., Generalized inverses of block triangular matrices, *SIAM J. Math. Appl.* 19(1970), 741-750.

[93] C. D. Meyer, Jr., Some remarks on EP_r matrices, and generalized inverses,*Linear Algebra and Appl.* 3 (1970), 275-278.

[94] C. D. Meyer, Jr., The Moore-Penrose inverse of a bordered matrix, *Linear Algebra Appl.* 5(1972), 375-382.

[95] C. D. Meyer, Jr., Generalized inverses and ranks of block matrices, *SIAM J. Appl. Math.* 25(1973), 597-602.

[96] J. Miao, The Moore-Penrose inverse of a rank modified matrix, *Numer. Math. J. Chinese Univ.* 11(1989), 355-361.

[97] J. Miao, General expression for the Moore-Penrose inverse of a 2×2 block matrix, *Linear Algebra Appl.* 151(1990), 1-15.

[98] S. K. Mitra, Fixed rank solution of linear matrix equation, *Sankhyā Ser. A* 35(1972), 387–392.

[99] S. K. Mitra, Common solutions to a pair of linear matrix equations $A_1XB_1 = C_1$ and $A_1XB_2 = C_2$, *Proc. Cambridge Philos. Soc.* 74(1973), 213-216.

[100] S. K. Mitra, Properties of the fundamental bordered matrix used in linear estimation, in Statistics and Probability, Essays in honor of C. R. Rao(G. Kallianpur et al, Eds), North Holland, New York, 1982.

[101] S. K. Mitra, The matrix equations $AX = C,\ XB = D$, *Linear Algebra Appl.* 59(1984), 171-181.

[102] S. K. Mitra, The minus partial order and the shorted matrix,*Linear Algebra Appl.* 83(1986), 1-27.

[103] S. K. Mitra, A pair of simultaneous linear matrix equations $A_1XB_1 = C_1$ and $A_2XB_2 = C_2$ and a programming problem, *Linear Algebra Appl.* 131(1990), 107-123.

[104] S. K. Mitra, Noncore square matrices miscellany, *Linear Algebra Appl.* 249(1996), 47–66.

[105] S. K. Mitra and R. E. Hartwig, Partial orderings based on the outer inverses, *Linear Algebra Appl.* 176(1992), 3–20.

[106] S. K. Mitra and M. L. Puri, Shorted matrices—an extended concept and some applications, *Linear Algebra Appl.* 42(1982), 57-79.

[107] S. K. Mitra and P. L. Odell, On parallel summability of matrices, Linear Algebra Appl. 74(1986), 239-255.

[108] M. Z. Nashed and X. Chen, Convergence of Newton-like methods for singular operator equations using outer inverses, *Numer. Math.* 66(1993), 235-257.

[109] D. V. Ouelette, Schur complements and statistics, *Linear Algebra Appl.* 36(1990), 187-295.

[110] A. B. Özgüler, The matrix equation $AXB + CYD = E$ over a principal ideal domain, *SIAM J. Matrix. Anal. Appl.* 12(1991), 581-591.

[111] A. B. Özgüler and N. Akar, A common solution to a pair of linear matrix equations over a principal ideal domain, *Linear Algebra Appl.* 144(1991), 85-99.

[112] W. V. Parker, The matrix equation $AX = XB$, *Duke Math. J.* 17(1950), 43-51.

[113] M. Pearl, On normal and EP_r matrices, *Michigan Math. J.* 6(1959), 1–5.

[114] M. Pearl, On normal EPr matrices, *Michigan Math. J.* 8(1961), 33-37.

[115] R. Penrose, A generalized inverse for matrices, *Proc. Cambridge Philos. Soc.* 51(1955), 406-413.

[116] L. Pernebo, An algebraic theory for design of controllers for linear multivariable systems—Part II: Feedback realizations and feedback design, *IEEE Trans. Automat. Control* 26(1981), 183-193.

[117] S. Puntanen, P. Šemrl and G. P. H. Styan, Some remarks on the parallel sum of two matrices. In Proceedings of the A. C. Aitken Centenary Conference, Dunedin, 1995, pp. 243-256.

[118] C. R. Rao and S. K. Mitra, Generalized Inverse of Matrices and Its Applications, Wiley, New York, 1971.

[119] C. R. Rao and H. Yanai, Generalized inverses of partitioned matrices useful in statistical applications, *Linear Algebra Appl.* 70(1985), 103-113.

[120] D. W. Robinson, Nullities of submatrices of the Moore-Penrose inverse, *Linear Algebra Appl.* 94(1987), 127-132.

[121] R. E. Roth, The equations $AX - YB = C$ and $AX - XB = C$ in matrices, *Proc. Amer. Math. Soc. A* 3(1952), 392-396.

[122] S. R. Searle, On inverting circulant matrices, *Linear Algebra Appl.* 25(1979), 77-89.

[123] N. Shinozaki and M. Sibuya, The reverse order law $(AB)^- = B^-A^-$, *Linear Algebra Appl.* 9(1974), 29-40.

[124] N. Shinozaki and M. Sibuya, Further results on the reverse order law, *Linear Algebra Appl.* 27(1979), 9-16.

[125] S. Slavova, G. Borisova and Z. Zelev, On the operator equation $AX = XB$. Complex analysis and applications '87 (Varna, 1987), 474-476, Bulgar. Acad. Sci., Sofia, 1989.

[126] R. L. Smith, Moore-Penrose inverses of block circulant and block k-circulant matrices, *Linear Algebra Appl.* 16(1979), 237-245.

[127] R. L. Smith, The Moore-Penrose inverse of a retrocirculant, *Linear Algebra Appl.* 22(1978), 1-8.

[128] G. P. H. Styan, Schur complements and linear models. Proc. First International Tampere Seminar on Linear Statistical Models and Their Applications, Dept. of Mathematical Science, University of Tampere, Finland, 1985, pp. 37-75.

[129] G. P. H. Styan and A. Takemura, Rank additivity and matrix polynomials. Studies in econometrics, time series, and multivariate statistics, 545–558, Academic Press, New York-London, 1983, pp. 545-558.

[130] W. Sun and Y. Wei, Inverse order rule for weighted generalized inverse, *SIAM J. Matrix Anal. Appl.* 19(1998), 72-775.

[131] Y. Tian, The general solution of the matrix equation $AXB = CYD$, *Math. Practice Theory* no. 1(1988), 61-63.

[132] Y. Tian, The Moore-Penrose inverse of a partitioned matrix and its applications, *Appl. Chinese Math. J. Chinese University* 7(1992), 310-314.

[133] Y. Tian, The Moore-Penrose inverse of a triple matrix product, *Math. in Theory and Practice* 1(1992), 64-70.

[134] Y. Tian, Calculating formulas for the ranks of submatrices in the Moore-Penrose inverse of a matrix, *J. Beijing Polytechnical University* 18(1992), 84-92.
[135] Y. Tian, Reverse order laws for the generalized inverses of multiple matrix products, *Linear Algebra Appl.* 211(1994), 185–200.
[136] Y. Tian, The Moore-Penrose inverses of $m \times n$ block matrices and their applications, *Linear Algebra Appl.* 283(1998), 35-60.
[137] Y. Tian, Universal similarity factorization equalities over real Clifford algebras, *Adv. Appl. Clifford algebras* 8(1998), 365–402.
[138] Y. Tian, Universal similarity factorization equalities over complex Clifford algebras, Proceedings of the 5th International Conference on Clifford Algebras, Ixtapa 1999, to appear.
[139] Y. Tian, The minimal rank of the matrix expression $A - BX - YC$, *Missouri J. Math. Sci.*, to appear.
[140] Y. Tian, The dimension of the intersection of k subspaces, *Missouri J. Math. Sci.*, to appear.
[141] G. E. Trapp, Inverses of circulant matrices and block circulant matrices, Kyungpook Math. J. 13(1973), 11-20.
[142] F. Uhlig, On the matrix equation $AX = B$ with applications to the generators of controllability matrix, *Linear Algebra Appl.* 85(1987), 203-209.
[143] J. Van der Woude, Feedback Decoupling and Stabilization for Linear System with Multiple Exogenous Variables, Ph. D. Thesis, Technical Univ. of Eindhoven, Netherlands, 1987.
[144] K. Wang, Generalizations of circulants, *Linear Algebra Appl.* 25(1979), 219-239.
[145] H. J. Waterhouse, Circulant style matrices closed under multiplication, *Linear and Multilinear Algebra* 18(1985), 197-206.
[146] M. Wei, Equivalent conditions for generalized inverses of products, *Linear Algebra Appl.* 266(1997), 347-363.
[147] Y. Wei, A characterization and representation of the generalized inverse $A^{(2)}_{S,T}$ and its application, *Linear Algebra Appl.* 280(1998), 87-86.
[148] Y. Wei and G. Wang, A survey on the generalized inverse $A^{(2)}_{S,T}$. In Proceedings of Meeting on Matrix Analysis and Applications, Spain, 1997, pp. 421-428.
[149] H. Werner, When is B^-A^- a generalized inverse of AB? *Linear Algebra Appl.* 210(1994), 255-263.
[150] H. Werner, G-inverses of matrix products. In Data Analysis and Statistics Inference (S. Schach and G. Trenkler, eds.), Eul-Verlag, Bergisch-Gladbach, 1992, pp. 531-546.
[151] E. A. Wibker, R. B. Howe, and J. D. Gilbert, Explicit solution to the reverse order law $(AB)^+ = B^-_{mr}A^-_{lr}$, *Linear Algebra Appl.* 25(1979), 107-114.
[152] H. J. Woerdeman, The lower order of triangular operators and minimal rank extensions, *Integral Equations operator Theory* 10(1987), 859-879.
[153] H. J. Woerdeman, Minimal rank completions for block matrices, *Linear Algebra Appl.* 121(1989), 105-122.
[154] H. J. Woerdeman, Minimal rank completions of partial banded matrices, *Linear and Multilinear Algebra* 36(1993), 59-69.
[155] G. Xu, M. Wei and D. Zhang, On solutions of matrix equation $AXB + CYD = F$, *Linear Algebra and Appl.* 279(1998), 93-109.